

CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
CC are also used in diagnosing a pathological condition or susceptibility  
CC to a pathological condition. The antibodies to the antigens can also  
CC be used in alleviating symptoms associated with the disorders and in  
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
CC immunoassays (ELISA). Disorders which are diagnosed or treated  
CC include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
CC and ocular disorders e.g. corneal infection. The polypeptides can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. Numerous  
CC examples of diseases and disorders treated by the nucleic acids and  
CC proteins are given in the specification. The present sequence

Query Match 92.5%; Score 37; DB 22; Length 33;  
Best Local Similarity 83.3%; Pred. No. 13;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
DB 9 adwtwh 14

RESULT 2  
AA06332  
ID AAY06332 standard; Protein; 103 AA.  
XX  
AC AAY06332;  
XX  
XX  
DT 06-SEP-1999 (first entry)  
XX  
XX  
DE Gliocladium roseum EgIII-like cellulase (partial sequence).  
XX  
XX  
KM Cellulase; endoglucanase; EgIII; textile; feed additive; baking;  
KM food processing; grain wet milling; pulp; paper.  
XX  
OS Gliocladium roseum.  
XX  
PN WO931255-A2.  
XX  
PD 24-JUN-1999.  
XX  
PF 14-DEC-1998; 98WO-US26552.  
XX  
PR 16-DEC-1997; 97US-0991720.  
XX  
PA (GEMV) GENENCOR INT INC.  
XX  
PI Bower BS, Fowler T, Phillips JT;  
XX  
DR WPI; 1999-395187/33.  
XX  
XX  
PT EgIII like cellulase  
XX  
PS Example; Fig 3; 47p; English.  
XX  
XX  
CC The present polypeptide represents a partial sequence of a novel  
CC EgIII-like cellulase of Gliocladium roseum. It was deduced from  
CC a partial gene sequence isolated from genomic DNA using PCR  
CC primers (see AAX59180-91) based on conserved motifs (see AAY06325-29)  
CC of Trichoderma reesei EgIII cellulase and related enzymes. PCR  
CC has been used to identify novel EgIII-like enzymes, including the  
CC present polypeptide, from bacterial and fungal sources (see  
CC AAY06331-70). Also provided by the invention are vectors, host

CC cells and methods for the recombinant production of such enzymes,  
CC which can be used in the treatment of cellulose-containing textiles,  
CC as feed additives, in the treatment of wood pulp, in the reduction  
CC of biomass to glucose, in the stone washing of indigo dyed denim,  
CC or as laundry detergent components (all claimed).  
XX  
SO Sequence 103 AA;

Query Match 92.5%; Score 37; DB 20; Length 103;  
Best Local Similarity 83.3%; Pred. No. 44;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
DB 29 adwsws 34

RESULT 3  
AA06363  
ID AAY06363 standard; Protein; 236 AA.  
XX  
AC AAY06363;  
XX  
XX  
DT 06-SEP-1999 (first entry)  
XX  
XX  
DE Gliocladium roseum EgIII-like cellulase.  
XX  
XX  
KM Cellulase; endoglucanase; EgIII; textile; feed additive; baking;  
KM food processing; grain wet milling; pulp; paper.  
XX  
XX  
OS Gliocladium roseum.  
XX  
PN WO931255-A2.  
XX  
PD 24-JUN-1999.  
XX  
PF 14-DEC-1998; 98WO-US26552.  
XX  
PR 16-DEC-1997; 97US-0991720.  
XX  
PA (GEMV) GENENCOR INT INC.  
XX  
PI Bower BS, Fowler T, Phillips JT;  
XX  
DR WPI; 1999-395187/33.  
XX  
XX  
PT EgIII like cellulase  
XX  
PS Example; Fig 6; 47p; English.  
XX  
XX  
CC The present polypeptide represents a full-length sequence of a  
CC novel EgIII-like cellulase of Gliocladium roseum. It was deduced  
CC from a gene sequence isolated from genomic DNA using PCR  
CC primers (see AAX59180-91) based on conserved motifs (see AAY06325-29)  
CC of Trichoderma reesei EgIII cellulase and related enzymes. PCR  
CC has been used to identify novel EgIII-like enzymes, including the  
CC present protein, from bacterial and fungal sources (see AAY06331-70).  
CC The sequence shows homology to T. reesei EgIII (see AAY06330). Also  
CC provided by the invention are vectors, host cells and methods  
CC for the recombinant production of such enzymes, which can be used  
CC in the treatment of cellulose-containing textiles, as feed  
CC additives, in the treatment of wood pulp, in the reduction of  
CC biomass to glucose, in the stone washing of indigo dyed denim, or  
CC as laundry detergent components (all claimed).  
XX  
SO Sequence 236 AA;

Query Match 92.5%; Score 37; DB 20; Length 236;  
Best Local Similarity 83.3%; Pred. No. 11e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2002, 11:50:50 ; Search time 10.4 Seconds

(without alignments)  
22.338 Million cell updates/sec

Title: US-09-643-260-6

Sequence: 1 ADMSWA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Database: SWISSProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	90.0	470	1	NRAM_IADBU
2	36	90.0	470	1	NRAM_IADBU
3	36	90.0	470	1	NRAM_IADBU
4	36	90.0	470	1	NRAM_IADBU
5	36	90.0	470	1	NRAM_IADBU
6	36	90.0	470	1	NRAM_IADBU
7	36	90.0	470	1	NRAM_IADBU
8	36	90.0	470	1	NRAM_IADBU
9	36	90.0	470	1	NRAM_IADBU
10	36	90.0	470	1	NRAM_IADBU
11	36	90.0	470	1	NRAM_IADBU
12	36	90.0	470	1	NRAM_IADBU
13	36	90.0	470	1	NRAM_IADBU
14	36	90.0	470	1	NRAM_IADBU
15	36	90.0	470	1	NRAM_IADBU
16	36	90.0	470	1	NRAM_IADBU
17	36	90.0	470	1	NRAM_IADBU
18	36	90.0	470	1	NRAM_IADBU
19	36	90.0	470	1	NRAM_IADBU
20	36	90.0	470	1	NRAM_IADBU
21	36	90.0	470	1	NRAM_IADBU
22	36	90.0	470	1	NRAM_IADBU
23	36	90.0	470	1	NRAM_IADBU
24	36	90.0	470	1	NRAM_IADBU
25	36	90.0	470	1	NRAM_IADBU
26	36	90.0	470	1	NRAM_IADBU
27	36	90.0	470	1	NRAM_IADBU
28	36	90.0	470	1	NRAM_IADBU
29	36	90.0	470	1	NRAM_IADBU
30	36	90.0	470	1	NRAM_IADBU
31	36	90.0	470	1	NRAM_IADBU
32	36	90.0	470	1	NRAM_IADBU
33	36	90.0	470	1	NRAM_IADBU

34	32	80.0	272	1	CYL_RHORI	P23135 rhodospirill
35	32	80.0	282	1	3ME2_ECOLI	P04395 escherichia
36	32	80.0	295	1	X769_HUMAN	Q09871 homo sapien
37	32	80.0	339	1	YJGB_ECOLI	P27250 escherichia
38	32	80.0	360	1	WNT2_CAEEL	P34889 caenorhabdi
39	32	80.0	362	1	DCUP_YEAST	P32347 saccharomyc
40	32	80.0	411	1	FOLC_BUCAI	P57265 buchnera ap
41	32	80.0	418	1	HUT_VIBPA	Q09289 vibrio para
42	32	80.0	453	1	NRAM_IAMIL	P03470 influenza a
43	32	80.0	454	1	NRAM_IAPUE	P03468 influenza a
44	32	80.0	470	1	NOS2_ONCMY	Q92091 oncorhynchu
45	32	80.0	536	1	YC42_SYNY3	P42349 synechocyst

## ALIGNMENTS

RESULT	ID	STANDARD	PRT	470 AA
NRAM_IADBU	NRAM_IADBU	NRAM_IADBU	NRAM_IADBU	NRAM_IADBU
AC	Q07570	NRAM_IADBU	NRAM_IADBU	NRAM_IADBU
DT	01-FEB-1995 (Rel. 31, last sequence update)			
DT	01-FEB-1995 (Rel. 31, last sequence update)			
DT	01-FEB-1996 (Rel. 33, last annotation update)			
DE	Neuraminidase (EC 3.2.1.18).			
GN	NA.			
OS	Influenza A virus (strain A/Duck/Burjatta/652/88).			
OC	Viruses: ssRNA negative-strand viruses: Orthomyxoviridae;			
OC	Influenza virus A and B group: Influenza A viruses;			
OC	Influenza A virus.			
OX	NCBI_TaxID=38956;			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=93212520; PubMed=8460490;			
RA	Saito T., Kawachi Y., Webster R.G.;			
RT	"Phylogenetic analysis of the N8 neuraminidase gene of influenza A			
RT	viruses."			
RL	Virology 193:868-876(1993).			
CC	- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE			
CC	FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO			
CC	AND FROM THE SITE OF INFECTION.			
CC	- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL			
CC	NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.			
CC	- SUBUNIT: HOMOTETRAMER.			
CC	- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED			
CC	SPIKE ON THE SURFACE OF THE VIRION			
CC	- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL: L06572; AAA43365.1; -			
DR	HSSP: P06820; 2BAT.			
DR	InterPro: IPR001860; Neur.			
DR	Pfam: PF00064; neur. 1.			
DR	ProDom: PD000431; Neur; 1.			
KW	Hydrolase; Glycosylase; Glycoprotein; Transmembrane.			
FT	DOMAIN 39..88			
FT	DOMAIN 89..470			
FT	ACT_SITE 273..273			
FT	ACT_SITE 275..275			
FT	CARBOHYD 46..46			
FT	CARBOHYD 54..54			
FT	CARBOHYD 144..144			
FT	CARBOHYD 293..293			
FT	CARBOHYD 398..398			
SO	SOURCE 470 AA; 51989 MW; DIA6F07460F8AD CRC64;			

Query Match 90.0%; Score 36; DB 1; Length 470;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
11111  
Db 453 ADMSW 457

RESULT 2  
NRAM\_IADCH STANDARD: PRT: 470 AA.  
ID NRAM\_IADCH

AC 007571; 275  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Neuraminidase (EC 3.2.1.18).  
GN NA.  
OS Influenza A virus (strain A/Duck/Chadovsk/1610/72).  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group; Influenza A viruses;  
OC Influenza A virus.  
OX NCBI\_TaxID=38957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93212520; PubMed=8460490;  
RA Saito T., Kawoka Y., Webster R.G.;  
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A viruses."  
RT Virology 193:868-876(1993).  
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO AND FROM THE SITE OF INFECTION.  
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED SPIKE ON THE SURFACE OF THE VIRION.  
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.  
CC CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: L06573; AAA43367.1; -.  
DR HSSP: P06820; 2BAT.  
DR InterPro: IPR001860; Neur.  
DR Pfam: PF00064; neur. 1.  
DR ProDom: PD000431; Neur. 1.  
KM Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.  
FT TRANSSEM 7 38 ANCHOR (BY SIMILARITY).  
FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.  
FT ACT\_SITE 89 470 HEAD OF NEURAMINIDASE.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
FT ACT\_SITE 275 275 BY SIMILARITY.  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 470 AA; 52070 MW; 169AB89FBE8006DC CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
11111  
Db 453 ADMSW 457

RESULT 3  
NRAM\_IADCH2 STANDARD: PRT: 470 AA.  
ID NRAM\_IADCH2

AC 007572; 275  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Neuraminidase (EC 3.2.1.18).  
GN NA.  
OS Influenza A virus (strain A/Duck/Hokkaido/8/80).  
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
OC Influenza virus A and B group; Influenza A viruses;  
OC Influenza A virus.  
OX NCBI\_TaxID=11358;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93212520; PubMed=8460490;  
RA Saito T., Kawoka Y., Webster R.G.;  
RT "Phylogenetic analysis of the N8 neuraminidase gene of influenza A viruses."  
RT Virology 193:868-876(1993).  
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO AND FROM THE SITE OF INFECTION.  
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED SPIKE ON THE SURFACE OF THE VIRION.  
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.  
CC CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: L06574; AAA43372.1; -.  
DR HSSP: P06820; 2BAT.  
DR InterPro: IPR001860; Neur.  
DR Pfam: PF00064; neur. 1.  
DR ProDom: PD000431; Neur. 1.  
KM Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.  
FT TRANSSEM 7 38 ANCHOR (BY SIMILARITY).  
FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.  
FT ACT\_SITE 89 470 HEAD OF NEURAMINIDASE.  
FT ACT\_SITE 273 273 BY SIMILARITY.  
FT ACT\_SITE 275 275 BY SIMILARITY.  
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 470 AA; 52015 MW; E1CID3E2C650B93C CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
11111  
Db 453 ADMSW 457

```

RESULT 4
NRAM_IADW2
ID NRAM_IADW2 STANDARD; PRT; 470 AA.
AC 007573;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Duck/Memphis/928/74).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11367;
RN [1]
RP MEDLINE=93212520; PubMed=8460490;
RX Saito T., Kawoka Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of influenza A
RT viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
CC FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
CC AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06575; AAA43404.1; -.
DR HSSP; P06820; 2B8T.
DR InterPro; IPR001860; Neur.
DR Pfam; PF00064; Neur. 1.
DR ProDom; PD000431; Neur. 1.
KW Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
DR TRANSMEM 7 38 ANCHOR (BY SIMILARITY).
FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.
FT ACT_SITE 89 273 HEAD OF NEURAMINIDASE.
FT ACT_SITE 273 275 BY SIMILARITY.
FT ACT_SITE 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 470 AA; 52146 MW; 30F5F9F364C1F49 CRC64;

```

Query Match 90.0%; Score 36; DB 1; Length 470;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
 DB 453 ADMSW 457

RESULT 5  
 NRAM\_IADU3 STANDARD; PRT; 470 AA.

```

DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Duck/Ukraine/1/63).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=11374;
RN [1]
RP MEDLINE=93212520; PubMed=8460490;
RX Saito T., Kawoka Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of influenza A
RT viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
CC FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
CC AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06576; AAA16234.1; -.
DR HSSP; P06820; 2B8T.
DR InterPro; IPR001860; Neur.
DR Pfam; PF00064; Neur. 1.
DR ProDom; PD000431; Neur. 1.
KW Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
DR TRANSMEM 7 37 ANCHOR (BY SIMILARITY).
FT DOMAIN 38 88 HYPERVARIABLE STALK REGION.
FT ACT_SITE 89 273 HEAD OF NEURAMINIDASE.
FT ACT_SITE 273 275 PROBABLE.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 470 AA; 51960 MW; B46D54A03AC84CCE CRC64;

```

Query Match 90.0%; Score 36; DB 1; Length 470;  
 Best Local Similarity 100.0%; Pred. No. 54;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
 DB 453 ADMSW 457

RESULT 6  
 NRAM\_IAGFN STANDARD; PRT; 470 AA.  
 ID NRAM\_IAGFN  
 AC 007574;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Neuraminidase (EC 3.2.1.18).  
 GN NA.  
 OS Influenza A virus (strain A/Guinea fowl/New York/4-3587/84).

```

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=38963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93212520; PubMed=8460490;
RA Saito T., Kawakita Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A
  viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
  FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
  AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
  NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
  SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L06584; AAA43428.1; -
DR HSSP: P06820; 2BRT.
DR InterPro: IPR001860; Neur.
DR Pfam: PF00064; neur; 1.
DR ProDom: PD000431; Neur; 1.
KM Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
FT TRANSMEM 7 38 ANCHOR (BY SIMILARITY).
FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.
FT DOMAIN 89 470 HEAD OF NEURAMINIDASE.
FT ACT_SITE 273 273 BY SIMILARITY.
FT ACT_SITE 275 275 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 470 AA; 52348 MW; D3BD2AAC0159PE66 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
DB 453 ADMSW 457

RESULT 7
NRAM_IAGHD STANDARD: PRT; 470 AA.
AC 007577;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Herring gull/DE/677/88).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OX NCBI_TaxID=38964;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=93212520; PubMed=8460490;
RA Saito T., Kawakita Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A
  viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
  FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
  AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
  NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
  SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L06585; AAA43368.1; -
DR HSSP: P06820; 2BRT.
DR InterPro: IPR001860; Neur.
DR Pfam: PF00064; neur; 1.
DR ProDom: PD000431; Neur; 1.
KM Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
FT TRANSMEM 7 38 ANCHOR (BY SIMILARITY).
FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.
FT DOMAIN 89 470 HEAD OF NEURAMINIDASE.
FT ACT_SITE 273 273 BY SIMILARITY.
FT ACT_SITE 275 275 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 470 AA; 52265 MW; 28AF0B75E80539B7 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
DB 453 ADMSW 457

RESULT 8
NRAM_IAGHD STANDARD: PRT; 470 AA.
AC 007578;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Egypt/311/89).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OX NCBI_TaxID=11401;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93212520; PubMed=8460490;
RA Saito T., Kawakita Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A
  viruses."

```

```

RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
CC FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
CC AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L06579; AAA43374.1; -.
DR HSSP: P06820; 2BAT.
DR InterPro: IPR001860; Neur.
DR Pfam: PF00064; neur; 1.
DR Prodom: PD000431; Neur; 1.
KW Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
FT TRANSMEM 7 38
FT DOMAIN 39 88
FT ACT_SITE 273 273 ANCHOR (BY SIMILARITY).
FT ACT_SITE 275 275 HYPERVARIABLE STALK REGION.
FT CARBOHYD 46 46 BY SIMILARITY.
FT CARBOHYD 54 54 BY SIMILARITY.
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 470 AA; 52234 MW; C50B21050A37668 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
DB 453 ADMSW 457

RESULT 9
NRAM_IATKL STANDARD; PRT; 470 AA.
ID NRAM_IATKL
AC 007583;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Mallard/Edmonton/2220/90).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=38965;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93212520; PubMed=8460490;
RA Saito T., Kawakita Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A
RT viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
CC FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
CC AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.

```

```

CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L06586; AAA43369.1; -.
DR HSSP: P06820; 2BAT.
DR InterPro: IPR001860; Neur.
DR Pfam: PF00064; neur; 1.
DR Prodom: PD000431; Neur; 1.
KW Hydroxylase; Glycosidase; Glycoprotein; Transmembrane.
FT TRANSMEM 7 38
FT DOMAIN 39 88
FT ACT_SITE 273 273 ANCHOR (BY SIMILARITY).
FT ACT_SITE 275 275 HYPERVARIABLE STALK REGION.
FT CARBOHYD 46 46 BY SIMILARITY.
FT CARBOHYD 54 54 BY SIMILARITY.
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 470 AA; 52070 MW; 557630C3E11F2765 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
DB 453 ADMSW 457

RESULT 10
NRAM_IATKL STANDARD; PRT; 470 AA.
ID NRAM_IATKL
AC 007585;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neuraminidase (EC 3.2.1.18).
GN NA.
OS Influenza A virus (strain A/Turkey/Minesota/501/78).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses;
OC Influenza A virus.
OX NCBI_TaxID=38984;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93212520; PubMed=8460490;
RA Saito T., Kawakita Y., Webster R.G.;
RT "Phylogenetic analysis of the N8 neuraminidase gene of Influenza A
RT viruses."
RL Virology 193:868-876(1993).
CC -1- FUNCTION: PREVENT SELF-AGGREGATION BY REMOVING THE CARBOHYDRATE
CC FROM THE VIRAL ENVELOPE. FACILITATE THE MOBILITY OF THE VIRUS TO
CC AND FROM THE SITE OF INFECTION.
CC -1- CATALYTIC ACTIVITY: CLEAVE THE TERMINAL SIALIC ACID (N-ACETYL
CC NEURAMINIC ACID) FROM CARBOHYDRATE CHAINS IN GLYCOPROTEINS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: VIRAL MEMBRANE. FORMS A MUSHROOM-SHAPED
CC SPIKE ON THE SURFACE OF THE VIRION.
CC -1- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L06588; AAA43410.1; -.
CC DR HSSP; P06820; ZBAT.
CC DR InterPro; IPR001860; Neur.
CC Pfam; PF00064; neur. 1.
CC DR ProDom; PD000431; Neur. 1.
CC KM Hydrolase; Glycosidase; Glycoprotein; Transmembrane.
CC FT TRANSMEM 7 38 ANCHOR (BY SIMILARITY).
CC FT DOMAIN 39 88 HYPERVARIABLE STALK REGION.
CC FT ACT_SITE 89 470 HEAD OF NEURAMINIDASE.
CC FT ACT_SITE 273 273 BY SIMILARITY.
CC FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 470 AA; 52352 MW; D573742ABF1E6B CRC64;

Query Match 90.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
Db 453 ADMSW 457

RESULT 11
MRJ5_APIME STANDARD; PRT; 598 AA.
ID MRJ5_APIME
AC 097432;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major royal jelly protein 5 precursor (MRJP-5) (Bee-milk protein).
GN MRJP5.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Head;
RX MEDLINE=99373663; PubMed=10441680;
RA Albert S.; Bhattacharya D.; Klaudiny J.; Schmitzova J.; Simuth J.;
RT "The family of major royal jelly proteins and its evolution.";
RL J. Mol. Evol. 49:290-297(1999).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN HONEYBEE NUTRITION. IT IS
CC FOUND IN THE ROYAL JELLY WHICH IS THE FOOD OF THE QUEEN HONEY BEE
CC LARVA. THE ROYAL JELLY DETERMINES THE DEVELOPMENT OF THE YOUNG
CC LARVAE AND IS RESPONSIBLE FOR THE HIGH REPRODUCTIVE ABILITY OF THE
CC HONEYBEE QUEEN.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: HYPOPHARYNGEAL GLANDS OF NURSE HONEY BEES.
CC -1- DEVELOPMENTAL STAGE: PRODUCED BY THE CEPHALIC GLANDULAR SYSTEM OF
CC THE NURSE HONEY BEE.
CC -1- SIMILARITY: BELONGS TO THE MAJOR ROYAL JELLY PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF004842; AAD01205.1; -.
CC DR InterPro; IPR003534; RoyalJelly.
CC DR Pfam; PR03022; MRJP. 1.
CC DR PRINTS; PR01366; ROYALJELLY.
CC FT SIGNAL 1 17 POTENTIAL.
CC FT CHAIN 18 598 MAJOR ROYAL JELLY PROTEIN 5.
CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 598 AA; 70236 MW; 2C603C77E7ACDF63 CRC64;

Query Match 90.0%; Score 36; DB 1; Length 598;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DMSWA 6
Db 113 DMSWA 117

RESULT 12
PGLR_PENGR STANDARD; PRT; 376 AA.
ID PGLR_PENGR
AC 093883;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Polylacturonase precursor (EC 3.2.1.15) (Pc) (Pectinase).
GN PG31.
OS Penicillium griseoeseum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=84562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCT 6421;
RA Ribon A.B.; Coelho J.L.C.; Barros E.G.; Araujo E.F.;
RT "Cloning and characterization of a gene encoding the
RT endopolylacturonase of Penicillium griseoeseum.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-
CC galactosiduronic linkages in pectate and other galacturonans.
CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
CC (POLYGLACTURONASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF085238; AAC83692.1; -.
CC DR InterPro; IPR000743; Polylacturonase.
CC DR Pfam; PF00295; Glyco_hydro_28; 1.
CC DR PROSITE; PS00502; POLYGLACTURONASE; 1.
CC KM Hydrolase; Glycosidase; Cell wall; Signal.
CC FT SIGNAL 1 20 POTENTIAL.
CC FT CHAIN 21 376 POLYGLACTURONASE.
CC SO SEQUENCE 376 AA; 38068 MW; 1EDB1EC56ED56928 CRC64;

Query Match 85.0%; Score 34; DB 1; Length 376;
Best Local Similarity 66.7%; Pred. No. 89;

```

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6

Db 349 SDMSWS 354

RESULT 13

FAST2\_SCHPO STANDARD; PRT; 1842 AA.

AC Q10289; P78973; 014163;

DT 15-JUL-1998 (Rel. 36, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fatty acid synthase subunit alpha (EC 2.3.1.86) [includes: Acyl carrier; 3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100) (beta-ketoacyl reductase); 3-oxoacyl-[acyl-carrier protein] synthase (EC 2.3.1.41) (beta-ketoacyl synthase)].

GN FAS2 OR LSD1 OR SPAC48.11C.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

OC NCBI\_TaxID=4896;

OX NCBI\_TaxID=4896;

RP SEQUENCE FROM N.A.

RA MEDLINE=96354912; PubMed=8769419;

RA Saitoh S., Takahashi K., Nabeshima K., Yamashita Y., Nakaseko Y., Hirata A., Yanagida M.;

RT "Aberrant mitosis in fission yeast mutants defective in fatty acid synthetase and acetyl CoA carboxylase.";

RT J. Cell Biol. 134:949-961(1996).

RL [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=21363051; PubMed=11470243;

RA Yokoyama K., Saitoh S., Ishida M., Yamakawa Y., Nakamura K., Inoue K., Taguchi R., Tokumura A., Nishijima M., Yanagida M., Setaka M.;

RT "Very long-chain fatty acid-containing phospholipids accumulate in fatty acid synthase temperature-sensitive mutant strains of the fission yeast Schizosaccharomyces pombe fas2/lsd1.";

RT Biochim. Biophys. Acta 1532:223-233(2001).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;

RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.

RN [4]

RP SEQUENCE OF 1-215 FROM N.A.

RA Koken M.H.M., de Rooij J.;

RL Submitted (Feb-1997) to the EMBL/Genbank/DBJ databases.

RN [5]

RP SEQUENCE OF 1-20.

RA MEDLINE=94245730; PubMed=8188691;

RA Kaeslin E., Heyer W.-D.;

RT "Schizosaccharomyces pombe fatty acid synthase mediates DNA strand exchange in vitro.";

RT J. Biol. Chem. 269:14103-14110(1994).

RL [6]

CC FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF THE ALPHA SUBUNIT CONTAINS DOMAINS FOR: ACYL CARRIER PROTEIN, 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND 3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE. THIS SUBUNIT COORDINATES THE BINDING OF THE SIX BETA SUBUNITS TO THE ENZYME COMPLEX.

CC CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH -> LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+).

CC CATALYTIC ACTIVITY: ACYL-[acyl-carrier protein] + malonyl-[acyl-carrier protein] -> 3-oxoacyl-[acyl-carrier protein] + CO(2) + [acyl-carrier protein].

CC CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] + NADP(+) -> 3-oxoacyl-[acyl-carrier protein] + NADPH.

CC SUBUNIT: FATTY ACID SYNTHETASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).

CC SIMILARITY: TO THE FATTY ACID SYNTHETASE, SUBUNIT ALPHA FROM

OTHER FUNGI.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: D83412; BA11913.1; -

DR EMBL: AB013747; BAB62029.1; -

DR EMBL: Z98762; CAB1481.1; -

DR EMBL: U82216; AAB39943.1; -

DR InterPro: IPR002582; ACPS.

DR InterPro: IPR000794; ketoacyl-synt.

DR InterPro: IPR003880; Phosphopant\_attach.

DR Pfam: PF01648; ACPS; 1.

DR Pfam: PF00109; ketoacyl-synt; 1.

DR Pfam: PF02801; ketoacyl-synt\_C; 1.

DR ProDom: PD004282; ACPS; 1.

DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 1.

DR PROSITE: PS00606; B\_KETOACYL\_SYNTHASE; 1.

KW Fatty acid biosynthesis; Multifunctional enzyme; Oxidoreductase; Transferase; NADP; Phosphopantetheine.

KW DOMAIN 1

FT DOMAIN 1 ?

FT DOMAIN 2 ?

FT BINDING 180 180 BETA-KETOACYL REDUCTASE.

FT ACT\_SITE 1262 1262 BETA-KETOACYL SYNTHASE (BY SIMILARITY).

FT CONFLICT 107 107 S -> A (IN REF. 4).

FT CONFLICT 422 422 K -> R (IN REF. 1).

SO SEQUENCE 1842 AA; 202168 MW; E4019F2D133EE571 CRC64;

Query Match 85.0%; Score 34; DB 1; Length 1842;

Best Local Similarity 66.7%; Pred. No. 3.8e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6

Db 400 SDMSWS 405

RESULT 14

NOS3\_SHEEP STANDARD; PRT; 99 AA.

AC P79209;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Nitric oxide synthase, endothelial (EC 1.14.13.39) (EC-NOS) (NOS, type IIT) (NOSIIT) (Endothelial NOS) (Constitutive NOS) (CNOS) (Fragment).

DE NOS3 OR ENOS.

GN NOS3 OR ENOS.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID=9940;

RP SEQUENCE FROM N.A.

RC Tissue-endothelial cells;

RA Aguan K., Weiner C.P.;

RT "Effect of hypoxia on the microvasculature of developing fetal brain of sheep: a studies on the expression pattern of constitutive forms of nitric oxide synthase.";

RT Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.

RL [1]

CC FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS IMPLICATED IN VASCULAR SMOOTH MUSCLE RELAXATION THROUGH A GMP-MEDIATED SIGNAL TRANSDUCTION PATHWAY. NO MEDIATES VASCULAR ENDOTHELIAL GROWTH FACTOR (VEGF)-INDUCED ANGIOGENESIS IN CORONARY VESSELS AND PROMOTES BLOOD CLOTTING THROUGH THE ACTIVATION OF PLATELETS (BY



```

CC          SIMILARITY).
CC -1- CATALYTIC ACTIVITY: L-arginine + N NADPH + M O(2) -> citrulline +
CC      nitric oxide + N NADP(+).
CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES
CC      TETRAHYDROBIOPTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF
CC      THE ENZYME (BY SIMILARITY).
CC -1- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN (BY
CC      SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U76738; AAB40705.1; -.
CC HSSP; P29473; INSE.
CC InterPro: IPR004030; NOS.
CC Pfam; PF02896; NO_synthase; 1.
CC PROSITE; PS60001; NOS; PARTIAL.
CC Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Calcium-binding;
CC Heme; Multigene family.
CC NON_TER 1
CC NON_TER 99
CC SEQUENCE 99 AA; 11034 MW; 82C3C765557031DA CRC64;

Query Match      82.5%; Score 33; DB 1; Length 99;
Best Local Similarity 80.0%; Pred. No. 37;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
   |||:|
Db 65 ADMAW 69

RESULT 15
Y132_METJA STANDARD; PRT; 220 AA.
AC 057596;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0132.
GN MJ0132.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N. A.
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: TO M.JANNASCHII MJ1220 AND MJEC142.
CC -1- SIMILARITY: WITH TYPE I RESTRICTION SYSTEM ADENINE METHYLASES
CC      (M SOBDUNIT).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U67470; AAB98113.1; -.
CC TIGR; MJ0132; -.
CC InterPro: IPR003356; N6_DNA_mtae.
CC Pfam; PF02384; N6_Mtae; 1.
CC Hypothetical protein; Complete proteome.
CC SEQUENCE 220 AA; 25766 MW; 710DDAE4C7A47954 CRC64;

```

```

Query Match      82.5%; Score 33; DB 1; Length 220;
Best Local Similarity 80.0%; Pred. No. 77;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5
   |||:|
Db 33 ADMAW 37

```

Search completed: September 3, 2002, 11:52:49  
Job time: 119 sec

Check 1 of 2

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2002, 11:48:39 : Search time 13.03 seconds  
(without alignments)  
11.247 Million cell updates/sec

Title: US-09-643-260-6  
Perfect score: 40  
Sequence: 1 ADMSWA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	34	85.0	44	US-08-905-223-274
2	33	82.5	74	US-08-379-538-2
3	33	82.5	136	US-08-774-065-2
4	33	82.5	218	US-08-032-848C-10
5	33	82.5	218	US-08-438-870-10
6	33	82.5	218	US-08-169-948B-34
7	33	82.5	218	US-08-448-873-34
8	33	82.5	218	US-08-382-452D-34
9	33	82.5	218	US-09-216-295-1
10	33	82.5	232	US-09-146-770-1
11	33	82.5	234	US-08-032-848C-9
12	33	82.5	234	US-08-438-870-9
13	33	82.5	234	US-09-146-770-3
14	33	82.5	234	US-09-146-770-4
15	33	82.5	234	US-09-216-295-3
16	33	82.5	234	US-09-216-295-4
17	33	82.5	239	US-09-216-295-15
18	33	82.5	467	US-08-140-104A-2
19	33	82.5	1205	US-07-908-245-2
20	33	82.5	1205	US-08-319-866-10
21	33	82.5	1205	US-09-123-708-6
22	33	82.5	1205	US-09-123-708-6
23	32	80.0	5	5217869-75
24	32	80.0	100	US-08-241-853-28
25	32	80.0	100	US-08-241-853-28
26	32	80.0	100	US-08-850-917-28
27	32	80.0	100	US-08-850-917-28

28	32	80.0	120	US-07-942-245-35	Sequence 35, Appl
29	32	80.0	170	US-09-199-637A-339	Sequence 339, Appl
30	32	80.0	260	US-09-216-295-23	Sequence 23, Appl
31	32	80.0	537	US-09-655-270A-11	Sequence 11, Appl
32	32	80.0	537	US-09-651-941-11	Sequence 11, Appl
33	32	80.0	616	US-09-136-574A-47	Sequence 47, Appl
34	32	80.0	677	US-08-836-567-2	Sequence 47, Appl
35	32	80.0	745	US-08-887-518-3	Sequence 3, Appl
36	32	80.0	745	US-09-023-321-3	Sequence 3, Appl
37	32	80.0	745	US-08-890-853-4	Sequence 3, Appl
38	32	80.0	745	US-09-032-475-3	Sequence 4, Appl
39	32	80.0	745	US-09-099-125A-4	Sequence 4, Appl
40	32	80.0	745	US-09-099-125A-4	Sequence 4, Appl
41	32	80.0	745	US-09-032-476-4	Sequence 4, Appl
42	32	80.0	745	US-08-890-854-4	Sequence 4, Appl
43	32	80.0	745	US-09-023-324-4	Sequence 4, Appl
44	32	80.0	745	US-09-168-629-2	Sequence 2, Appl
45	32	80.0	745	US-08-910-820-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-08-905-223-274  
Sequence 274, Application US/08905223  
Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Duclert, Aymeric  
APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5 ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klobbe, Martens, Olsson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Neda A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-0176  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: -26...-1  
IDENTIFICATION METHOD: Von Heljine matrix  
OTHER INFORMATION: score 9.6  
OTHER INFORMATION: seq WLTALASWALC/RI  
US-08-905-223-274

Query Match 85.0%; Score 34; DB 4; Length 44;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
11111  
Db 19 ASMSWA 24

## RESULT 2

US-08-379-538-2  
; Sequence 2, Application US/08379538  
; Patent No. 5804554  
; GENERAL INFORMATION:  
; APPLICANT: Volkman, Robert A.  
; APPLICANT: Saccomano, Nicholas A.  
; APPLICANT: Nason II, Deane M.  
; APPLICANT: Heck, Steven D.  
; APPLICANT: Ronau, Robert T.  
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES  
; TITLE OF INVENTION: FROM FILISTATA HIBERNALIS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pfizer Inc  
; STREET: 235 East 42nd Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/379,538  
; FILING DATE: 3-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/887073  
; FILING DATE: 21-MAY-1992  
; APPLICATION NUMBER: PCT/US93/03921  
; FILING DATE: 30-APRIL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ziellinski, Bryan  
; REGISTRATION NUMBER: 34,462  
; REFERENCE/DOCKET NUMBER: PC8175A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 573-4585  
; TELEFAX: (212) 573-1939  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Filistata hibernalis  
; TISSUE TYPE: venom  
; US-08-379-538-2

Query Match 82.5%; Score 33; DB 1; Length 74;  
Best Local Similarity 80.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DMSWA 6  
11111  
Db 52 DMSWS 56

RESULT 3  
US-08-774-065-2  
; Sequence 2, Application US/08774065  
; Patent No. 5989899  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Benjamin  
; APPLICANT: Clarkson, Kathleen  
; APPLICANT: Larenas, Edmund  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: NOVEL OVERSIZED CELLULOSE COMPOSITIONS  
; TITLE OF INVENTION: FOR USE IN DETERGENT COMPOSITIONS AND  
; TITLE OF INVENTION: IN THE TREATMENT OF TEXTILES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENENCOR INTERNATIONAL  
; STREET: 925 PAGE MILL ROAD  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/774,065  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gialster, Debra J.  
; REGISTRATION NUMBER: 33,888  
; REFERENCE/DOCKET NUMBER: GC368  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-846-7620  
; TELEFAX: 415-845-6504  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-774-065-2

Query Match 82.5%; Score 33; DB 2; Length 136;  
Best Local Similarity 66.7%; Pred. No. 11+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
11111  
Db 62 ADQWS 67

RESULT 4  
US-08-032-848C-10  
; Sequence 10, Application US/08032848C  
; Patent No. 5475101  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Michael  
; APPLICANT: Clarkson, Kathleen A.  
; APPLICANT: Weiss, Geoffrey L.  
; APPLICANT: Larenas, Edward  
; APPLICANT: Lorch, Jeffrey D.  
; TITLE OF INVENTION: Purification and Molecular Cloning of  
; TITLE OF INVENTION: EG III Cellulase  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International

STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/032,848C  
FILING DATE: MAR 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.  
REGISTRATION NUMBER: 33,401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7356  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-032-848C-10

Query Match 82.5%; Score 33; DB 1; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 5  
US-08-438-870-10  
Sequence 10, Application US/08438870  
Patent No. 5753484  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Latenas, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of EG  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,870  
FILING DATE: May 10, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Christopher L. Stone  
REGISTRATION NUMBER: 35,696  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7555

TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-870-10

Query Match 82.5%; Score 33; DB 1; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 6  
US-08-169-948B-34  
Sequence 34, Application US/08169948B  
Patent No. 5861271  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen  
APPLICANT: Collier, Katherine  
APPLICANT: Latenas, Edmund  
TITLE OF INVENTION: No. 5861271el Cellulase Enzymes and Systems  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/169,948B  
FILING DATE: DEC 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.  
REGISTRATION NUMBER: 33,401  
REFERENCE/DOCKET NUMBER: GC226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7536  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-169-948B-34

Query Match 82.5%; Score 33; DB 2; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 7  
US-08-448-873-34  
; Sequence 34, Application US/08448873  
; Patent No. 5874276  
; GENERAL INFORMATION:  
; APPLICANT: Fowler, Timothy  
; APPLICANT: Clark, Michael  
; APPLICANT: Clarkson, Kathleen  
; APPLICANT: Collier, Katherine A.  
; APPLICANT: Larenas, Edmund  
; TITLE OF INVENTION: No. 5874276 Cellulase Enzymes and Systems  
; TITLE OF INVENTION: For their Expressions  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International  
; STREET: 180 Kimball Way  
; CITY: South San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,873  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/169,948  
; FILING DATE: 17-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stone, Christopher L.  
; REGISTRATION NUMBER: 35,696  
; REFERENCE/DOCKET NUMBER: GC226D14  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 742-7555  
; TELEFAX: (415) 742-7217  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 218 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-448-873-34

Query Match 82.5%; Score 33; DB 2; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMWSA 6  
||| |  
Db 46 ADMOWS 51

RESULT 8  
US-08-382-452D-34  
; Sequence 34, Application US/08382452D  
; Patent No. 6268196  
; GENERAL INFORMATION:  
; APPLICANT: Fowler, Timothy  
; APPLICANT: Clarkson, Kathleen A.  
; APPLICANT: Ward, Michael  
; APPLICANT: Collier, Katherine D.  
; APPLICANT: Larenas, Edmund A.  
; TITLE OF INVENTION: NOVEL CELLULOSE ENZYMES AND SYSTEMS  
; TITLE OF INVENTION: FOR THEIR EXPRESSION  
; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genencor International  
; STREET: 925 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/382,452D  
; FILING DATE: February 1, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Christopher L. Stone  
; REGISTRATION NUMBER: 36,696  
; REFERENCE/DOCKET NUMBER: GC226-2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 742-7555  
; TELEFAX: (415) 742-7217  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 218 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-382-452D-34

Query Match 82.5%; Score 33; DB 4; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMWSA 6  
||| |  
Db 46 ADMOWS 51

RESULT 9  
US-09-216-295-1  
; Sequence 1, Application US/09216295  
; Patent No. 6268328  
; GENERAL INFORMATION:  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: Wendt, Dan J.  
; TITLE OF INVENTION: No. 6268328 Variant EGIII-Like Cellulase Compositions  
; FILE REFERENCE: GC555  
; CURRENT APPLICATION NUMBER: US/09/216,295  
; CURRENT FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO: 1  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Trichoderma longibrachiatum  
; US-09-216-295-1

Query Match 82.5%; Score 33; DB 4; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMWSA 6  
||| |  
Db 46 ADMOWS 51

RESULT 10  
US-09-146-770-1  
; Sequence 1, Application US/09146770

Patent No. 6187732  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
TITLE OF INVENTION: Mutant EgIII Cellulase, DNA Encoding  
FILE REFERENCE: GC546  
CURRENT APPLICATION NUMBER: US/09/146,770  
CURRENT FILING DATE: 1998-09-03  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 232  
TYPE: PRT  
ORGANISM: T. reesei  
US-09-146-770-1

Query Match 82.5%; Score 33; DB 4; Length 232;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADWSMA 6  
1111:  
DB 60 ADWQMS 65

RESULT 11  
US-08-032-848C-9  
Sequence 9, Application US/08032848C  
Patent No. 5475101  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Larenas, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/032,848C  
FILING DATE: MAR 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.  
REGISTRATION NUMBER: 33,401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7356  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-032-848C-9

Query Match 82.5%; Score 33; DB 1; Length 234;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ADWSMA 6  
1111:  
DB 62 ADWQMS 67

RESULT 12  
US-08-438-870-9  
Sequence 9, Application US/08438870  
Patent No. 5753484  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Larenas, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of EG  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,870  
FILING DATE: May 10, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Christopher L. Stone  
REGISTRATION NUMBER: 35,696  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7555  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-870-9

Query Match 82.5%; Score 33; DB 1; Length 234;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADWSMA 6  
1111:  
DB 62 ADWQMS 67

RESULT 13  
US-09-146-770-3  
Sequence 3, Application US/09146770  
Patent No. 6187732  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
TITLE OF INVENTION: Mutant EgIII Cellulase, DNA Encoding  
FILE REFERENCE: GC546  
CURRENT APPLICATION NUMBER: US/09/146,770  
CURRENT FILING DATE: 1998-09-03  
NUMBER OF SEQ ID NOS: 4

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: T. reesei
US-09-146-770-3

```

Search completed: September 3, 2002, 11:51:39  
Job time: 180 sec

```

Query Match      82.5%; Score 33; DB 4; Length 234;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 ADMSWA 6
      ||| |
      62 ADMQMS 67

```

```

RESULT 14
US-09-146-770-4
; Sequence 4, Application US/09146770
; Patent No. 6187732
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; TITLE OF INVENTION: Mutant EgII Cellulase, DNA Encoding
; TITLE OF INVENTION: Such EgII Compositions and Methods for Obtaining Same
; FILE REFERENCE: GC546
; CURRENT APPLICATION NUMBER: US/09/146,770
; CURRENT FILING DATE: 1998-09-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 234
; TYPE: PRT
; ORGANISM: H. schweinitzii
US-09-146-770-4

```

```

Query Match      82.5%; Score 33; DB 4; Length 234;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 ADMSWA 6
      ||| |
      62 ADMQMS 67

```

```

RESULT 15
US-09-216-295-3
; Sequence 3, Application US/09216295
; Patent No. 6268328
; GENERAL INFORMATION:
; APPLICANT: Mitchinson, Colin
; APPLICANT: Mendt, Dan J.
; TITLE OF INVENTION: No. 6268328el Variant EgIII-Like Cellulase Compositions
; FILE REFERENCE: GC555
; CURRENT APPLICATION NUMBER: US/09/216,295
; CURRENT FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-09-216-295-3

```

```

Query Match      82.5%; Score 33; DB 4; Length 234;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 ADMSWA 6
      ||| |
      62 ADMQMS 67

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2002, 11:49:24 : Search time 14.83 seconds  
(without alignments)  
38.876 Million cell updates/sec

Title: US-09-643-260-6

Perfect score: 40

Sequence: 1 ADMSWA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: pirl:\*  
2: pirl:\*  
3: pirl:\*  
4: pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	92.5	610	2 T35222	hypothetical prote
2	37	92.5	885	2 C83441	two-component sens
3	36	90.0	374	2 B83241	conserved hypothet
4	36	90.0	889	2 E87304	non-dependent rec
5	34	85.0	275	2 T05822	hypothetical prote
6	34	85.0	617	2 C84922	hypothetical prote
7	34	85.0	1842	2 T43409	probable protein x
8	34	85.0	1842	2 T38781	probable fatty-acid
9	33	82.5	132	2 S65785	fatty acid synthas
10	33	82.5	198	2 B82531	mel-13a protein -
11	33	82.5	220	2 D64316	conserved hypothet
12	33	82.5	232	2 D58353	restriction modifi
13	33	82.5	239	2 D90470	CD1b protein - she
14	33	82.5	257	2 D87152	hypothetical prote
15	33	82.5	276	2 B75337	conserved hypothet
16	33	82.5	304	2 F83632	hypothetical prote
17	33	82.5	324	2 AB3548	probable cytochrom
18	33	82.5	328	2 E72424	vegetalidile Incomp
19	33	82.5	333	2 S47246	oligopeptide ABC t
20	33	82.5	350	2 S71923	gene CDI protein -
21	33	82.5	368	2 H90998	cysteine proteinas
22	33	82.5	410	2 D75475	probable proteinas
23	33	82.5	415	2 AE1844	lycopene cyclase -
24	33	82.5	418	2 AE1460	hypothetical prote
25	33	82.5	418	2 AF1097	sugar ABC transpor
26	33	82.5	421	2 T38242	probable phosphata
27	33	82.5	433	2 T31511	hypothetical prote
28	33	82.5	467	2 G82697	hypothetical prote
29	33	82.5	478	2 E89790	6-phospho-beta-glu

#### ALIGNMENTS

RESULT 1  
T35222  
hypothetical protein SC5C7.15 SC5C7.15 - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T35222  
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1998  
A:Reference number: Z21572  
A:Accession: T35222  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-610 <SEED>  
A:Cross-references: EMBL:AL031515; PDB:CAA20627.1; GSPDB:GN00070; SCOEDB:SC5C7.15  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC5C7.15

Query Match 92.5%; Score 37; DB 2; Length 610;

Best Local Similarity 83.3%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0;

OY 1 ADMSWA 6  
Db 83 ADMSWA 88

#### RESULT 2

C83441  
two-component sensor KdpD PA1636 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83441

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; )  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p  
A:Reference number: AB2950; MUID:20437337  
A:Accession: C83441  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-885 <STC>  
A:Cross-references: GB:AE004591; GB:AE004091; NID:99947599; PDB:AG05025.1; GSPDB:GI  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: kdpD; PA1636

Query Match 92.5%; Score 37; DB 2; Length 885;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;



Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ADMSWA 6  
 |||:|  
 Db 563 ADMWA 568

## RESULT 3

B83241

Conserved hypothetical protein PA3230 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: B83241

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: B83241

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 &lt;STO&gt;

A:Cross-references: GB:AE004746; GB:AE004091; NID:g9949350; PIDN:AAG0618.1; GSPDB:GN001

C:Genetics:

A:Experimental source: strain PA01

A:Gene: PA3230

Query Match 90.0%; Score 36; DB 2; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 92;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DMSWA 6  
 |||:|  
 Db 81 DMSWA 85

## RESULT 4

E87304

TonB-dependent receptor [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: E87304

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: E87304

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-889 &lt;STO&gt;

A:Cross-references: GB:AE005673; NID:g13421615; PIDN:AAK22433.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0446

Query Match 90.0%; Score 36; DB 2; Length 889;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 |||:|  
 Db 618 ADMSW 622

## RESULT 5

T05822

hypothetical protein T5K18.170 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 13-Aug-1999

C:Accession: T05822

R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.;

submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15453

A:Accession: T05822

A:Molecule type: DNA

A:Residues: 1-275 &lt;BEV&gt;

A:Cross-references: EMBL:AL022580

A:Experimental source: cultivar Columbia; BAC clone T5K18

C:Genetics:

A:Map position: 4

A:Introns: 103/3; 141/3; 169/1; 206/3

A:Note: T5K18.170

Query Match 85.0%; Score 34; DB 2; Length 275;  
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
 |||:|  
 Db 57 SDMSWS 62

## RESULT 6

C84922

probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: C84922

R:Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Omayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84922

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:g4249408; PIDN:AAD13705.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g48010

A:Map position: 2

Query Match 85.0%; Score 34; DB 2; Length 617;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
 |||:|  
 Db 500 ADMWAS 505

## RESULT 7

T43409

probable fatty-acid synthase (EC 2.3.1.85) alpha chain - fission yeast (Schizosacchar

N:Alternate names: fatty acid synthetase alpha subunit

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000

C:Accession: T43409

R:Saltooh, S.; Takahashi, K.; Nabeshima, K.; Yamashita, Y.; Nakaseko, Y.; Hirata, A.;

J. Cell Biol. 134, 949-961, 1996

A:Title: Aberrant mitosis in fission yeast mutants defective in fatty acid synthetase

A:Reference number: Z22493; MUID:96354912

A:Accession: T43409

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1842 &lt;SAT&gt;

A:Cross-references: EMBL:D83412; NID:g1199959; PIDN:BA11913.1; PID:g1199960

C:Genetics:

A:Note: ISd1+

C:Superfamily: yeast fatty-acid synthase  
C:Keywords: acyltransferase; coenzyme A

Query Match 85.0%; Score 34; DB 2; Length 1842;  
Best Local Similarity 66.7%; Pred. No. 9.4e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
DB 400 SDWNMA 405

RESULT 8  
T38781  
Fatty acid synthase, subunit alpha - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T38781  
R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21751  
A:Accession: T38781  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-1842 <SKE>  
A:Cross-references: EMBL:Z98762; PIDN:CAH1481.1; GSPDB:GN00066; SPDB:SPAC4A8.11c  
A:Experimental source: strain 972H-; cosmid c4A8  
C:Genetics:  
A:Gene: SPDB:SPAC4A8.11c  
A:Map position: 1  
C:Superfamily: yeast fatty-acid synthase

Query Match 85.0%; Score 34; DB 2; Length 1842;  
Best Local Similarity 66.7%; Pred. No. 9.4e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
DB 400 SDWNMA 405

RESULT 9  
S65785  
mel-13a protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 26-Aug-1999  
C:Accession: S65785  
R:Yatsu, O.; Kanno, R.; Isono, K.; Taniguchi, M.; Kanno, M.  
Biochem. Biophys. Acta 1305, 109-112, 1996  
A:Title: Cloning and characterization of two transcripts generated from the mel-13 gene  
A:Reference number: S65785; MUID:96180310  
A:Accession: S65785  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-132 <TET>  
A:Cross-references: EMBL:U35309  
C:Genetics:  
A:Gene: mel-13  
C:Superfamily: mouse mel-13a protein  
C:Keywords: alternative splicing

Query Match 82.5%; Score 33; DB 2; Length 132;  
Best Local Similarity 80.0%; Pred. No. 96;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
DB 57 SDMSW 61

RESULT 10  
B82531  
Conserved hypothetical protein XF2666 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: B82531  
R:anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: B82531; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82531  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-198 <SIM>  
A:Cross-references: GB:AE004072; GB:AE003849; NID:g9107884; PIDN:AAFR5463.1; GSPDB:GI  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Bridges, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carriaro, D.M.; Carre  
as-Melo, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; F  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Saw  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2666  
C:Superfamily: conserved hypothetical protein MUI677

Query Match 82.5%; Score 33; DB 2; Length 198;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DMSWA 6  
DB 135 DWNMA 139

RESULT 11  
D64316  
restriction modification enzyme subunit M1 homolog - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: D64316  
R:Balt, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blai  
Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannas  
A:Reference number: A64300; MUID:96337999  
A:Accession: D64316  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-220 <BDL>  
A:Cross-references: GB:U67470; GB:L77117; NID:g2826247; PIDN:AA98113.1; PID:g1592267  
C:Genetics:  
A:Map position: REV127472-126810  
A:Start codon: TTG

Query Match 82.5%; Score 33; DB 2; Length 220;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5

Db 33 ADMAM 37

# RESULT 12

S58353

CD1b protein - sheep (fragment)

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 14-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 21-Jan-2000

C:Accession: S58353

R:Ferguson, E.D.; Dutla, B.M.; Hein, W.; Hopkins, J.

submitted to the EMBL Data Library, July 1995

A:Description: The ovine CD1 gene family contains at least four CD1b homologues.

A:Reference number: S58353

A:Accession: S58353

A:Status: preliminary

A:Molecule type: mRNA

A:Cross-references: EMBL:X90567; NID:g945010; PIDN:CAA62187.1; PID:g945011

C:Superfamily: class I histocompatibility antigen, immunoglobulin homology

F:115-180/Domain: Immunoglobulin homology <IMM>

Query Match 82.5%; Score 33; DB 2; Length 232;  
Best Local Similarity 80.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

# RESULT 13

D90470

hypothetical protein cysH [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001

C:Accession: D90470

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: D90470

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-239 <KUR>

A:Cross-references: GB:AE006641; NID:g13816282; PIDN:AAK43019.1; GSPDB:GN00155

C:Genetics:

A:Gene: cysH

C:Superfamily: 3'-phosphoadenosine 5'-phosphosulfate reductase

Query Match 82.5%; Score 33; DB 2; Length 239;  
Best Local Similarity 80.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

# RESULT 14

D87152

conserved hypothetical protein ML1945 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C:Accession: D87152

R:Coile, S.T.; Biglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho-  
eam, M.A.; Rutherford, K.M.

Net#re 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;

A:Title: Massive gene decay in the leprosy bacillus

A:Reference number: A86909; PMID:21128732; PMID:11234002

A:Accession: D87152

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-257 <STO>

A:Cross-references: GB:AL450380; NID:g13093601; PIDN:CAC30900.1; GSPDB:GN00147

C:Genetics:

A:Gene: ML1945

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1100

Query Match 82.5%; Score 33; DB 2; Length 257;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

# RESULT 15

B75337

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: B75337

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; PMID:20036896

A:Accession: B75337

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-276 <NHT>

A:Cross-references: GB:AE002032; GB:AE000513; NID:g6459715; PIDN:AAF1479.1; PID:g6645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1923

A:Map position: 1

Query Match 82.5%; Score 33; DB 2; Length 276;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Search completed: September 3, 2002, 11:52:02  
Job time: 158 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:47:19 ; Search time 13.02 Seconds  
(without alignments)  
11.256 Million cell updates/sec

Title: US-09-643-260-6  
Perfect score: 40  
Sequence: 1 ADMSWA 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCPUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	85.0	44	4	US-08-905-223-274 Sequence 274, App
2	33	82.5	74	1	US-08-379-538-2 Sequence 2, Appli
3	33	82.5	136	2	US-08-774-065-2 Sequence 2, Appli
4	33	82.5	218	1	US-08-032-848C-10 Sequence 10, Appli
5	33	82.5	218	1	US-08-438-870-10 Sequence 10, Appli
6	33	82.5	218	2	US-08-169-948B-34 Sequence 34, Appli
7	33	82.5	218	4	US-08-448-873-34 Sequence 34, Appli
8	33	82.5	218	4	US-08-382-452D-34 Sequence 34, Appli
9	33	82.5	218	4	US-09-216-295-1 Sequence 1, Appli
10	33	82.5	222	4	US-09-146-770-1 Sequence 1, Appli
11	33	82.5	234	1	US-08-032-848C-9 Sequence 9, Appli
12	33	82.5	234	1	US-08-438-870-9 Sequence 9, Appli
13	33	82.5	234	4	US-09-146-770-3 Sequence 3, Appli
14	33	82.5	234	4	US-09-146-770-4 Sequence 4, Appli
15	33	82.5	234	4	US-09-216-295-3 Sequence 3, Appli
16	33	82.5	234	4	US-09-216-295-4 Sequence 4, Appli
17	33	82.5	239	4	US-09-216-295-15 Sequence 15, Appli
18	33	82.5	467	1	US-08-140-104A-2 Sequence 2, Appli
19	33	82.5	1205	1	US-07-908-245-10 Sequence 10, Appli
20	33	82.5	1205	2	US-08-319-866-6 Sequence 6, Appli
21	33	82.5	1205	4	US-09-123-708-6 Sequence 10, Appli
22	33	82.5	1205	4	US-09-123-708-6 Sequence 6, Appli
23	33	82.5	1205	4	US-09-123-708-6 Sequence 6, Appli
24	32	80.0	100	1	US-08-241-853-28 Sequence 29, Appli
25	32	80.0	100	1	US-08-241-853-28 Sequence 29, Appli
26	32	80.0	100	2	US-08-850-917-28 Sequence 29, Appli
27	32	80.0	100	2	US-08-850-917-28 Sequence 29, Appli

28	32	80.0	120	1	US-07-942-245-35 Sequence 35, Appli
29	32	80.0	170	4	US-09-199-637A-339 Sequence 339, App
30	32	80.0	260	4	US-09-216-295-23 Sequence 23, Appli
31	32	80.0	537	4	US-09-655-270A-11 Sequence 11, Appli
32	32	80.0	537	4	US-09-651-941-11 Sequence 11, Appli
33	32	80.0	616	4	US-09-136-574A-47 Sequence 2, Appli
34	32	80.0	677	4	US-08-836-567-2 Sequence 2, Appli
35	32	80.0	745	2	US-08-887-518-3 Sequence 3, Appli
36	32	80.0	745	2	US-09-023-321-3 Sequence 3, Appli
37	32	80.0	745	2	US-08-890-853-4 Sequence 4, Appli
38	32	80.0	745	2	US-09-032-475-3 Sequence 3, Appli
39	32	80.0	745	2	US-09-099-125A-4 Sequence 4, Appli
40	32	80.0	745	2	US-09-099-124A-4 Sequence 4, Appli
41	32	80.0	745	4	US-09-032-476-4 Sequence 4, Appli
42	32	80.0	745	4	US-08-890-854-4 Sequence 4, Appli
43	32	80.0	745	4	US-09-023-324-4 Sequence 4, Appli
44	32	80.0	745	4	US-09-168-629-2 Sequence 2, Appli
45	32	80.0	745	4	US-08-910-820-10 Sequence 10, Appli

#### ALIGNMENTS

RESULT 1  
US-08-905-223-274  
Sequence 274, Application US/08905223  
Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Duclert, Aymeric  
APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESS: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: -26...-1  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 9.6  
OTHER INFORMATION: seq WLIALASWALC/RI  
US-08-905-223-274

Query Match 85.0%; Score 34; DB 4; Length 44;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMWSA 6  
11111  
DB 19 ASMSWA 24

## RESULT 2

US-08-379-538-2  
; Sequence 2, Application US/08379538  
; Patent No. 5804554  
; GENERAL INFORMATION:  
; APPLICANT: Volkman, Robert A.  
; APPLICANT: Saccomano, Nicholas A.  
; APPLICANT: Nason II, Deane M.  
; APPLICANT: Heck, Steven D.  
; APPLICANT: Ronau, Robert T.  
; TITLE OF INVENTION: CALCIUM CHANNEL BLOCKING POLYPEPTIDES  
; TITLE OF INVENTION: FROM FILISTARA HIBERNALIS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pfizer Inc  
; STREET: 235 East 42nd Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10017  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/379,538  
; FILING DATE: 3-MAY-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/887073  
; FILING DATE: 21-MAY-1992  
; APPLICATION NUMBER: PCT/US93/03921  
; FILING DATE: 30-APRIL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zielinski, Bryan  
; REGISTRATION NUMBER: 34,462  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 573-4585  
; TELEFAX: (212) 573-1939  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 74 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Filistara hibernalis  
; TISSUE TYPE: venom  
; US-08-379-538-2

Query Match 82.5%; Score 33; DB 1; Length 74;  
Best Local Similarity 80.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DMSWA 6  
11111  
DB 52 DMSWS 56

RESULT 3  
US-08-774-065-2  
; Sequence 2, Application US/08774065  
; Patent No. 5989899  
; GENERAL INFORMATION:  
; APPLICANT: Bower, Benjamin  
; APPLICANT: Clarkson, Kathleen  
; APPLICANT: Larens, Edmund  
; APPLICANT: Ward, Michael  
; TITLE OF INVENTION: NOVEL OVERSIZED CELLULOSE COMPOSITIONS  
; TITLE OF INVENTION: FOR USE IN DIFFERENT COMPOSITIONS AND  
; TITLE OF INVENTION: IN THE TREATMENT OF TEXTILES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENECOR INTERNATIONAL  
; STREET: 925 PAGE MILL ROAD  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/774,065  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Glaister, Debra J.  
; REGISTRATION NUMBER: 33,888  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-846-7620  
; TELEFAX: 415-846-6504  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 136 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-774-065-2

Query Match 82.5%; Score 33; DB 2; Length 136;  
Best Local Similarity 66.7%; Pred. No. 1,1e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMWSA 6  
11111  
DB 62 ADMOWS 67

RESULT 4  
US-08-032-848C-10  
; Sequence 10, Application US/08032848C  
; Patent No. 5475101  
; GENERAL INFORMATION:  
; APPLICANT: Ward, Michael  
; APPLICANT: Clarkson, Kathleen A.  
; APPLICANT: Weiss, Geoffrey L.  
; APPLICANT: Larens, Edward  
; APPLICANT: Lorch, Jeffrey D.  
; TITLE OF INVENTION: Purification and Molecular Cloning of  
; TITLE OF INVENTION: EG III Cellulase  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genecor International

STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/032,848C  
FILING DATE: MAR 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.  
REGISTRATION NUMBER: 33,401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7356  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-032-848C-10

Query Match 82.5%; Score 33; DB 1; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 5  
US-08-438-870-10  
Sequence 10, Application US/08438870  
Patent No. 5753484  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Larens, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of EG  
TITLE OF INVENTION: IIT Cellulase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,870  
FILING DATE: May 10, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Christopher L. Stone  
REGISTRATION NUMBER: 35,696  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7555

TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-870-10

Query Match 82.5%; Score 33; DB 1; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 6  
US-08-169-948B-34  
Sequence 34, Application US/08169948B  
Patent No. 5861271  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen  
APPLICANT: Collier, Katherine  
APPLICANT: Larens, Edmund  
TITLE OF INVENTION: No. 5861271el Cellulase Enzymes and Systems  
TITLE OF INVENTION: For Their Expression  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/169,948B  
FILING DATE: DEC 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Horn, Margaret A.  
REGISTRATION NUMBER: 33,401  
REFERENCE/DOCKET NUMBER: GC226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7536  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-169-948B-34

Query Match 82.5%; Score 33; DB 2; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ADMSWA 6  
||| |  
Db 46 ADMSWA 51

RESULT 7  
US-08-448-873-34  
Sequence 34, Application US/08448873  
Patent No. 5874276  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen  
APPLICANT: Collier, Katherine A.  
APPLICANT: Larenaas, Edmund  
TITLE OF INVENTION: No. 5874276el Cellulase Enzymes and Systems  
TITLE OF INVENTION: For Their Expressions  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,873  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/169,948  
FILING DATE: 17-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stone, Christopher L.  
REGISTRATION NUMBER: 35,696  
REFERENCE/DOCKET NUMBER: GC226D14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-448-873-34

Query Match 82.5%; Score 33; DB 2; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADM5MA 6  
DB 46 ADMQMS 51

RESULT 8  
US-08-382-452D-34  
Sequence 34, Application US/08382452D  
Patent No. 6268196  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Ward, Michael  
APPLICANT: Collier, Katherine D.  
APPLICANT: Larenaas, Edmund A.  
TITLE OF INVENTION: NOVEL CELLULOSE ENZYMES AND SYSTEMS  
TITLE OF INVENTION: FOR THEIR EXPRESSION  
NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 925 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,452D  
FILING DATE: February 1, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Christopher L. Stone  
REGISTRATION NUMBER: 36,696  
REFERENCE/DOCKET NUMBER: GC226-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 742-7555  
TELEFAX: (415) 742-7217  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-382-452D-34

Query Match 82.5%; Score 33; DB 4; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADM5MA 6  
DB 46 ADMQMS 51

RESULT 9  
US-09-216-295-1  
Sequence 1, Application US/09216295  
Patent No. 6268328  
GENERAL INFORMATION:  
APPLICANT: Mitchinson, Colin  
APPLICANT: Wendt, Dan J.  
TITLE OF INVENTION: No. 6268328el Variant EGIII-Like Cellulase Compositions  
FILE REFERENCE: GC555  
CURRENT APPLICATION NUMBER: US/09/216,295  
CURRENT FILING DATE: 1998-12-18  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 1  
LENGTH: 218  
TYPE: PRT  
ORGANISM: Trichoderma longibrachiatum  
US-09-216-295-1

Query Match 82.5%; Score 33; DB 4; Length 218;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADM5MA 6  
DB 46 ADMQMS 51

RESULT 10  
US-09-146-770-1  
Sequence 1, Application US/09146770

Patent No. 6187732  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
TITLE OF INVENTION: Mutant Egit Cellulase, DNA Encoding  
FILE REFERENCE: GC546  
CURRENT APPLICATION NUMBER: US/09/146,770  
CURRENT FILING DATE: 1998-09-03  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 232  
TYPE: PRT  
ORGANISM: T. reesei  
US-09-146-770-1

Query Match 82.5%; Score 33; DB 4; Length 232;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
Db 60 ADMSWA 65

RESULT 11  
US-08-032-848C-9  
Sequence 9, Application US/08032848C  
Patent No. 5475101  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Larenas, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of  
TITLE OF INVENTION: Egit Cellulase  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/032,848C  
FILING DATE: MAR 17 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hohn, Margaret A.  
REGISTRATION NUMBER: 33,401  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7356  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-032-848C-9

Query Match 82.5%; Score 33; DB 1; Length 234;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
Db 62 ADMSWA 67

RESULT 12  
US-08-438-870-9  
Sequence 9, Application US/08438870  
Patent No. 5753484  
GENERAL INFORMATION:  
APPLICANT: Ward, Michael  
APPLICANT: Clarkson, Kathleen A.  
APPLICANT: Weiss, Geoffrey L.  
APPLICANT: Larenas, Edward  
APPLICANT: Lorch, Jeffrey D.  
TITLE OF INVENTION: Purification and Molecular Cloning of Egit Cellulase  
TITLE OF INVENTION: Egit Cellulase  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International  
STREET: 180 Kimball Way  
CITY: South San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,870  
FILING DATE: May 10, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Christopher L. Stone  
REGISTRATION NUMBER: 35,696  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415 742-7555  
TELEFAX: 415 742-7217  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 234 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-870-9

Query Match 82.5%; Score 33; DB 1; Length 234;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
Db 62 ADMSWA 67

RESULT 13  
US-09-146-770-3  
Sequence 3, Application US/09146770  
Patent No. 6187732  
GENERAL INFORMATION:  
APPLICANT: Fowler, Timothy  
TITLE OF INVENTION: Mutant Egit Cellulase, DNA Encoding  
FILE REFERENCE: GC546  
CURRENT APPLICATION NUMBER: US/09/146,770  
CURRENT FILING DATE: 1998-09-03  
NUMBER OF SEQ ID NOS: 4



SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 234  
 ; TYPE: PRT  
 ; ORGANISM: T. reesel  
 US-09-146-770-3

Search completed: September 7, 2002, 10:51:33  
 Job time: 254 sec

Query Match  
 Best Local Similarity 82.5%; Score 33; DB 4; Length 234;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSMA 6  
 ||| |:  
 Db 62 ADMQWS 67

RESULT 14  
 US-09-146-770-4  
 ; Sequence 4, Application US/09146770  
 ; Patent No. 6187732  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fowler, Timothy  
 ; TITLE OF INVENTION: Mutant EgIII Cellulase, DNA Encoding  
 ; FILE REFERENCE: GC346  
 ; CURRENT APPLICATION NUMBER: US/09/146,770  
 ; CURRENT FILING DATE: 1998-09-03  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 234  
 ; TYPE: PRT  
 ; ORGANISM: H. schweinfeltzii  
 US-09-146-770-4

Query Match  
 Best Local Similarity 82.5%; Score 33; DB 4; Length 234;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSMA 6  
 ||| |:  
 Db 62 ADMQWS 67

RESULT 15  
 US-09-216-295-3  
 ; Sequence 3, Application US/09216295  
 ; Patent No. 6268328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mitchinson, Colin  
 ; APPLICANT: Wendt, Dan J.  
 ; TITLE OF INVENTION: No. 6268328el Variant EgIII-like Cellulase Compositions  
 ; FILE REFERENCE: GC555  
 ; CURRENT APPLICATION NUMBER: US/09/216,295  
 ; CURRENT FILING DATE: 1998-12-18  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 234  
 ; TYPE: PRT  
 ; ORGANISM: Trichoderma reesel  
 US-09-216-295-3

Query Match  
 Best Local Similarity 82.5%; Score 33; DB 4; Length 234;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ADMSMA 6  
 ||| |:  
 Db 62 ADMQWS 67

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 3, 2002, 11:50:24 ; Search time 25.22 Seconds  
(without alignments)  
41.157 Million cell updates/sec

Title: US-09-643-260-6  
Perfect score: 40  
Sequence: 1 ADMSWA 6

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	92.5	437	16	Q92K30 rhizobium m
2	37	92.5	548	16	Q92M15 rhizobium m
3	37	92.5	597	5	Q9VGP2 drosophila m
4	37	92.5	610	2	Q86712 streptomyces
5	37	92.5	885	16	Q91389 pseudomonas
6	36	90.0	205	2	Q9ACR5 streptomyces
7	36	90.0	242	12	Q919R8 culex nigri
8	36	90.0	358	10	Q50002 pinus arme
9	36	90.0	374	16	Q9H210 pseudomonas
10	36	90.0	452	4	Q96AB7 homo sapien
11	36	90.0	477	11	Q9CYU6 mus musculu
12	36	90.0	484	4	Q9BTV6 homo sapien
13	36	90.0	598	5	Q97432 apis mellif
14	36	90.0	889	16	Q9AAZ6 caulobacter
15	36	90.0	1005	10	Q9XG22 arabidopsis
16	36	90.0	5435	2	Q9LAX2 streptomyces

17	34	85.0	273	10	Q94JM4 arabidopsis
18	34	85.0	273	10	Q940D6 arabidopsis
19	34	85.0	275	10	Q65710 arabidopsis
20	34	85.0	376	3	Q9UVL4 penicillium
21	34	85.0	617	10	P93050 arabidopsis
22	34	85.0	1842	3	Q96PM8 schizosacch
23	34	85.0	1842	3	Q96WT7 schizosacch
24	34	85.0	1842	3	Q96WT6 schizosacch
25	34	85.0	98	5	Q9VB45 xylella fas
26	33	82.5	161	11	Q921P9 ratius norv
27	33	82.5	198	13	Q9PA54 meleagrid h
28	33	82.5	213	12	Q9ELI7 meleagrid h
29	33	82.5	213	12	Q9DPT1 meleagrid h
30	33	82.5	234	3	Q00095 trichoderma
31	33	82.5	239	17	Q97UT3 mycobacteri
32	33	82.5	257	16	Q9X787 mycobacteri
33	33	82.5	276	16	Q9RT43 delinococcus
34	33	82.5	304	16	Q91719 pseudomonas
35	33	82.5	309	2	Q9F163 amycolatops
36	33	82.5	316	2	Q69348 rhodococcus
37	33	82.5	320	4	Q96JY4 rhodococcus
38	33	82.5	323	6	Q9RT79 ovis aries
39	33	82.5	328	16	Q9WXR6 thermotoga
40	33	82.5	330	4	Q96WZ6 streptomyces
41	33	82.5	332	2	Q9RJM7 streptomyces
42	33	82.5	335	16	Q989F6 rhizobium 1
43	33	82.5	338	2	Q46645 erwinia amy
44	33	82.5	350	10	Q41057 pisum sativ
45	33	82.5	350	10	Q949A2 pisum sativ

## ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	437 AA.
Q92K30	Q92K30			
AC	Q92K30			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHEtical PROTEIN SMC01671.			
GN	SMC01671.			
OS	Rhizobium melioli (Sinorhizobium melioli).			
OC	Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group.			
OC	Rhizobiaceae, Sinorhizobium.			
NC	NCBI_TaxID=382;			
ON	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-1021:			
RX	MEDLINE-21368234; PubMed-11474104;			
RA	Galibert F., Finan T.W., Long S.R., Puehler A., Abola P., Ampe F.,			
RA	Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Botte G.,			
RA	Boutry M., Bowser L., Bumester J., Cadieu E., Capela D., Chain P.,			
RA	Cowie A., Davis R.W., Dreano S., Federpriel N.A., Fisher R.F.,			
RA	Glox S., Godrie T., Goffeau A., Golding B., Gouy J., Gurjal M.,			
RA	Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,			
RA	Kahn M.L., Kalmann S., Keating D.H., Kiss E., Komp C., Lelaure V.,			
RA	Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,			
RA	Ramsberger U., Surzycki R., Thepault P., Vandenbol M.,			
RA	Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.,			
RT	"The composite genome of the legume symbiont Sinorhizobium melioli";			
RL	Science 293:666-672(2001).			
KW	EMBL: AL591790; CAC46862.1; -			
DR	Hypothetical protein; Complete proteome.			
SO	SEQUENCE 437 AA; 48372 MW; 950E0B3DA963CE78 CRC64;			

Query Match 92.5%; Score 37; DB 16; Length 437;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE-97000351; PubMed-8843435;  
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL031513; CAA20627.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 610 AA; 67368 MW; 052CEA90DB589021 CRC64;

Query Match 92.5%; Score 37; DB 2; Length 610;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ADMSWA 6  
 DB 83 ADMAMA 88

RESULT 5  
 ID Q91389 PRELIMINARY; PRT; 885 AA.  
 AC Q91389;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE TWO-COMPONENT SENSOR KDP.  
 GN KDPD OR PA1636.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15692 / PA01;  
 RA MEDLINE-20437337; PubMed-10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey R.L., Britman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber M.J., Collier L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith J.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RT Nature 406:959-964(2000).  
 RL -1- SIMILARITY: TO: PROKARYOTE SENSORY TRANSDUCTION PROTEINS.  
 CC EMBL; AE004591; AAG05025.1; -;  
 DR HSSP; P02933; 1JOY.  
 DR InterPro: IPR004358; BCTRLSENSOR.  
 DR InterPro: IPR003018; GAF.  
 DR InterPro: IPR003594; HATPase\_C.  
 DR InterPro: IPR00361; His\_kina.  
 DR InterPro: IPR004359; His\_KIN\_sig.  
 DR InterPro: IPR003852; KdpD.  
 DR Pfam; PF02518; HATPase\_C; 1.  
 DR Pfam; PF02702; KdpD; 1.  
 DR Pfam; PF00512; signal; 1.  
 DR PRINTS; PR00344; BCTRLSENSOR.  
 DR SMART; SM00065; GAF; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00388; HisKA; 1.  
 KW Complete proteome; kinase; phosphorylation; sensory transduction;  
 KW Transference.  
 SQ SEQUENCE 885 AA; 97019 MW; 20FC8E2B2A876C0 CRC64;

Query Match 92.5%; Score 37; DB 16; Length 885;  
 Best Local Similarity 83.3%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
 DB 563 ADMAMA 568

RESULT 6  
 ID Q9ACR5 PRELIMINARY; PRT; 205 AA.  
 AC Q9ACR5;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE HYPOTHETICAL 23.1 KDA PROTEIN.  
 GN SCP1.253.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Brown S.P., Murphy L.D., Harris D.;  
 RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Bentley S.D., Parkhill J., Barrett B.G., Rajandream M.A.;  
 RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE-98241550; PubMed-9573173;  
 RA Redenbach M., Ikeda K., Yamasaki M., Kinashi H.;  
 RT "Cloning and physical mapping of the EcoRI fragments of the giant  
 RT linear plasmid SCP1.";  
 RL J. Bacteriol. 180:2796-2799(1998).  
 DR EMBL; AL590464; CAC36779.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 205 AA; 23051 MW; 6602396CF93F2D9 CRC64;

Query Match 90.0%; Score 36; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 DB 10 ADMWSW 14

RESULT 7  
 ID Q919K8 PRELIMINARY; PRT; 242 AA.  
 AC Q919K8;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CUN068 HYPOTHETICAL PROTEIN.  
 GN CUN068.  
 OS Culex nigripalpus baculovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
 OX NCBI\_TaxID=130556;  
 RN [1]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-FLORIDA1997;  
 RA MEDLINE-21488685; PubMed-11602755;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kutish G.F.;  
 RT "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus.";  
 RL J. Virol. 75:11157-11165(2001).  
 RN [2]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN-FLORIDA1997;

RA Alfonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.J., Rock D.L., Kutish G.F.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF403738; AAK94146.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 242 AA; 27222 MW; 6014967531110E52 CRC64;

Query Match 90.0%; Score 36; DB 12; Length 242;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DMSWA 6  
 |||||  
 DB 80 DMSWA 84

RESULT 8  
 ID 050002 PRELIMINARY; PRT; 358 AA.  
 AC 050002;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE CYSTEINE PROTEASE.  
 OS Prunus armeniaca (Apricot).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
 OC NCBI\_TaxID=36596;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERGERON: TISSUE-MESOCARP PLUS EXOCARP;  
 RA Mbeague-A-Mbeague D., Gomez R.-M., Fils-Lycaon B.;  
 RT "Sequence of Atp1, a Cysteine Proteinase From Apricot Fruit  
 (Accession No. U93166). Gene Expression During Fruit Ripening. (PGR97-  
 RT 179).";  
 RL Plant Physiol. 115:1730-1730(1997).  
 DR EMBL: U93166; AAB97142.1; -  
 DR HSP: P07711; ICLT.  
 DR MEROPS: C01.041; -  
 DR InterPro: IPR003015; HLH\_Myc.  
 DR InterPro: IPR000668; Peptidase\_C1.  
 DR InterPro: IPR000169; Thiolprot\_act\_site.  
 DR Pfam: PF00112; Peptidase\_C1; 1.  
 DR PRINTS: PR00705; PAPA1N.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 DR PROSITE: PS00640; THIOL\_PROTEASE\_ASN; 1.  
 DR PROSITE: PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW Hydrolase; Protease; Thiol protease.  
 SQ SEQUENCE 358 AA; 39309 MW; C98F78793B002554 CRC64;

Query Match 90.0%; Score 36; DB 10; Length 358;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
 |||||  
 DB 108 ADMSW 112

RESULT 9  
 ID 09H210 PRELIMINARY; PRT; 374 AA.  
 AC 09H210;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
 DE HYPOTHELICAL PROTEIN PA3230.  
 GN PA3230.  
 OS - Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OC NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15692 / PA01;  
 RA MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle N.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltz L., Tolerino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004746; AAG0618.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 374 AA; 42269 MW; 31EF185C4F683884 CRC64;

Query Match 90.0%; Score 36; DB 16; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DMSWA 6  
 |||||  
 DB 81 DMSWA 85

RESULT 10  
 ID 096AB7 PRELIMINARY; PRT; 452 AA.  
 AC 096AB7;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HYPOTHELICAL 50.6 KDA PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-MELANOMA;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC017335; AAH17335.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 452 AA; 50575 MW; B79D25EE38096733 CRC64;

Query Match 90.0%; Score 36; DB 4; Length 452;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
 |||||  
 DB 338 ADMSW 342

RESULT 11  
 ID 09CY06 PRELIMINARY; PRT; 477 AA.  
 AC 09CY06;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 2810443J12RIK PROTEIN.  
 GN 2810443J12RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glisic C., King B., Kocikawa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,  
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilmink L.,  
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -i SIMILARITY: CONTAINS 2 WD\_REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AK013297; BAB28775.1; -;  
 DR MGD; MGI:1914478; 2810443J12R1k.  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 2.  
 DR SMART; SM00320; WD40; 4.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 1.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Repeat; WD repeat.  
 SQ SEQUENCE 477 AA; 53201 MW; 26557352A4BA9C CRC64;

Query Match 90.0%; Score 36; DB 11; Length 477;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 Db 337 ADMSW 341

RESULT 12  
 Q9BTV6 PRELIMINARY; PRT; 484 AA.  
 AC Q9BTV6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE HYPOTHEMETICAL 54.1 KDA PROTEIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN; NEUROBLASTOMA;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 CC -i SIMILARITY: CONTAINS 2 WD\_REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; BC003123; AA03123.1; -;  
 DR InterPro: IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 2.  
 DR SMART; SM00320; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_2.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 1.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 KW Hypothetical protein; Repeat; WD repeat.

FT NON\_TER 1  
 SQ SEQUENCE 484 AA; 54088 MW; 1A2CA3237CB7358E CRC64;

Query Match 90.0%; Score 36; DB 4; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 Db 370 ADMSW 374

RESULT 13  
 ID 097432 PRELIMINARY; PRT; 598 AA.  
 AC 097432;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE MAJOR ROYAL JELLY PROTEIN MRP5 PRECURSOR.  
 GN MRP5.  
 OS Apis mellifera (Honeybee).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;  
 OC Apoidea; Apidae; Apis.  
 OX NCBI\_TaxID=7460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=NUCLEUS HEADS;  
 RX MEDLINE=99373663; PubMed-10441680;  
 RA Albert S., Bhattacharya D., Klaidiny J., Schmitzova J., Simth J.,  
 RT "The family of major royal jelly proteins and its evolution."  
 RL J. Mol. Evol. 49:280-297(1999).  
 DR EMBL; AF004642; AAD01205.1; -;  
 DR InterPro: IPR003534; RoyalJelly.  
 DR Pfam; PF03022; MRJP; 1.  
 DR PRINTS; PR01366; ROYALJELLY.  
 KW signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 SQ SEQUENCE 598 AA; 70236 MW; 2C603C77E7ACDF63 CRC64;

Query Match 90.0%; Score 36; DB 5; Length 598;  
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DMSWA 6  
 Db 113 DMSWA 117

RESULT 14  
 ID 09AAZ6 PRELIMINARY; PRT; 889 AA.  
 AC 09AAZ6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE TONB-DEPENDENT RECEPTOR.  
 GN CC0446.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed-11259647;  
 RA Nierman W.C., Felblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocak I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,

RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uterback T., Tran K., Wolf A., Yamathavan J., Erniolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of *Caulobacter crescentus*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL: AF005717; AAK22433.1; -.  
DR TIGR: CC0446; -.  
DR InterPro: IPR000531; TonB\_boxC.  
DR Pfam: PF00593; TonB\_boxC; 1.  
KW Receptor; Complete proteome.  
SO SEQUENCE 889 AA; 95775 MW; 75FCBD7A726A01A5 CRC64;

Query Match 90.0%; Score 36; DB 16; Length 889;  
Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
DB 618 ADMSW 622

## RESULT 15

O9XG22 PRELIMINARY; PRT; 1005 AA.  
ID O9XG22;  
AC O9XG22;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE T1N24.22 PROTEIN.  
GN T1N24.22.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Mashu;  
RT "The A. thaliana Genome Sequencing Project";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Murray J., Langston Y., Clarke K., Drone K.;  
RT "The sequence of A. thaliana T1N24.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Waterston R.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL: AF149413; AAD40144.1; -.  
DR InterPro: IPR000719; Euk\_Pkinase.  
DR InterPro: IPR001611; LRR.  
DR InterPro: IPR003592; LRR\_out.  
DR InterPro: IPR002290; Ser\_thr\_kinase.  
DR Pfam: PF00560; LRR; 19.  
DR Pfam: PF00069; Pkinase; 1.  
DR PRINTS: PRO0019; LEURICHRP.  
DR SMART: SM00370; LRR; 17.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Serine/threonine-protein kinase; transferase.  
KW SEQUENCE 1005 AA; 111963 MW; BB006438CC9541C9 CRC64;  
SO SEQUENCE

Query Match 90.0%; Score 36; DB 10; Length 1005;  
Best Local Similarity 100.0%; Pred. No. 6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
DB 906 ADMSW 910

Search completed: September 3, 2002, 11:52:34  
Job time: 130 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 3, 2002, 11:48:09 : Search time 29.94 Seconds  
(without alignments)  
22.259 Million cell updates/sec

Title: US-09-643-260-6  
Perfect score: 40  
Sequence: 1 ADMSMA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Genesec.032802:\*  
1: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	92.5	33	22	AAU21305
2	37	92.5	103	20	AAU06332
3	37	92.5	236	20	AAU06363
4	37	92.5	236	21	AAU1876
5	37	92.5	236	21	AAU84341
6	37	92.5	597	22	ABB62635
7	37	92.5	885	22	AAU33594
8	36	90.0	50	22	AAU65714
9	36	90.0	174	22	AAU65780
10	36	90.0	218	18	AAU30839
11	36	90.0	225	22	AAU65779

12	36	90.0	276	22	AAU39444
13	36	90.0	277	21	AAU42053
14	36	90.0	289	22	AAU24678
15	36	90.0	371	22	AAU41230
16	36	90.0	378	22	AAU65775
17	36	90.0	452	22	AAU3545
18	36	90.0	821	22	ABG22441
19	36	90.0	5435	22	AAU10145
20	34	85.0	44	20	AAU12990
21	34	85.0	67	22	AAU45676
22	34	85.0	86	20	AAU36020
23	34	85.0	166	22	AAU38883
24	34	85.0	336	22	AAU39139
25	34	85.0	351	22	AAU23394
26	33	82.5	74	14	AAU40036
27	33	82.5	84	22	AAU51313
28	33	82.5	98	22	ABU71225
29	33	82.5	103	20	AAU6339
30	33	82.5	136	19	AAU69222
31	33	82.5	147	22	ABU11725
32	33	82.5	216	21	AAU3400
33	33	82.5	218	15	AAU63169
34	33	82.5	227	22	AAU67460
35	33	82.5	232	20	AAU06330
36	33	82.5	232	21	AAU69498
37	33	82.5	232	21	AAU84327
38	33	82.5	234	15	AAU63168
39	33	82.5	234	16	AAU79541
40	33	82.5	234	16	AAU77265
41	33	82.5	234	17	AAU02034
42	33	82.5	234	20	AAU06350
43	33	82.5	234	21	AAU14862
44	33	82.5	234	21	AAU14863
45	33	82.5	234	21	AAU69499

#### ALIGNMENTS

RESULT 1  
AAU21305 standard; Protein; 33 AA.  
AC AAU21305;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Human novel foetal antigen, SEQ ID NO 1549.  
XX  
KW Human: foetal tissue antigen; antiinflammatory; neuroprotective;  
KW immunomodulator; cardiovascular; cytosolic; nephrotoxic;  
KW cardiovascular; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; breast neoplasm; cancer;  
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
KW cerebral ischemia; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; epithelial cell proliferation; food additive.  
XX  
OS Homo sapiens.  
XX  
PN MO200155312-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01321.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.



PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214866.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 03-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0234999.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488782/53.  
N-PSDB; AAS34125.

New polynucleotides and polypeptides for diagnosing, treating, preventing or prognosis e.g. diseases or disorders of the nervous, musculoskeletal, excretory, gastrointestinal, reproductive, and respiratory systems

Claim 11; SEQ ID No 1549; 642bp; English.

The invention relates to novel nucleic acids encoding novel human foetal antigens. The nucleic acids and proteins are used to prevent, treat (e.g.

CC by gene therapy) or ameliorate a medical condition in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They  
 CC are also used in diagnosing a pathological condition or susceptibility  
 CC to a pathological condition. The antibodies to the antigens can also  
 CC be used in alleviating symptoms associated with the disorders and in  
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated  
 CC include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
 CC and ocular disorders e.g. corneal infection. The polypeptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to  
 CC prevent skin aging due to sunburn, to maintain organs before  
 CC transplantation, for supporting cell culture of primary tissues, to  
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
 CC as a food additive or preservative to increase or decrease storage  
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
 CC minerals, cofactors and other nutritional components. Numerous  
 CC examples of diseases and disorders treated by the nucleic acids and  
 CC proteins are given in the specification. The present sequence

Query Match 92.5%; Score 37; DB 22; Length 33;  
 Best Local Similarity 83.3%; Pred. No. 13;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSMA 6  
 DB 9 adwtwa 14

RESULT 2  
 AAY06332  
 ID AAY06332 standard; Protein; 103 AA.

XX AAY06332;  
 XX  
 AC AAY06332;  
 XX  
 DT 06-SEP-1999 (first entry)  
 XX  
 DE Glucoladum roseum Egitr-like cellulase (partial sequence).  
 XX  
 KM Cellulase; endoglucanase; Egitr; textile; feed additive; baking;  
 KM food processing; grain wet milling; pulp; paper.  
 XX  
 OS Glucoladum roseum.  
 OS  
 PN WO9931255-A2.  
 XX  
 PD 24-JUN-1999.  
 PD  
 PF 14-DEC-1998; 98WO-US26552.  
 PF  
 FR 16-DEC-1997; 97US-0991720.  
 FR  
 PA (GENEV ) GENENCOR INT INC.  
 PA  
 PI Bower BS, Fowler T, Phillips JT;  
 PI  
 XX WPI; 1999-395187/33.  
 DR  
 XX Egitr like cellulase  
 PT  
 PS Example; Fig 3; 47pp; English.  
 XX  
 CC The present polypeptide represents a partial sequence of a novel  
 CC Egitr-like cellulase of Glucoladum roseum. It was deduced from  
 CC a partial gene sequence isolated from genomic DNA using PCR  
 CC primers (see AAX59180-91) based on conserved motifs (see AAY06325-29)  
 CC of Trichoderma reesei Egitr cellulase and related enzymes. PCR  
 CC has been used to identify novel Egitr-like enzymes, including the  
 CC present polypeptide, from bacterial and fungal sources (see  
 CC AAY06331-70). Also provided by the invention are vectors, host

CC cells and methods for the recombinant production of such enzymes,  
 CC which can be used in the treatment of cellulose-containing textiles,  
 CC as feed additives, in the treatment of wood pulp, in the reduction  
 CC of biomass to glucose, in the stone washing of indigo dyed denim,  
 CC or as laundry detergent components (all claimed).  
 CC  
 SQ Sequence 103 AA;

Query Match 92.5%; Score 37; DB 20; Length 103;  
 Best Local Similarity 83.3%; Pred. No. 44;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSMA 6  
 DB 29 adwaws.34

RESULT 3  
 AAY06363  
 ID AAY06363 standard; Protein; 236 AA.

XX AAY06363;  
 XX  
 AC AAY06363;  
 XX  
 DT 06-SEP-1999 (first entry)  
 XX  
 DE Glucoladum roseum Egitr-like cellulase.  
 XX  
 KM Cellulase; endoglucanase; Egitr; textile; feed additive; baking;  
 KM food processing; grain wet milling; pulp; paper.  
 XX  
 OS Glucoladum roseum.  
 OS  
 PN WO9931255-A2.  
 XX  
 PD 24-JUN-1999.  
 PD  
 PF 14-DEC-1998; 98WO-US26552.  
 PF  
 PR 16-DEC-1997; 97US-0991720.  
 PR  
 PA (GENEV ) GENENCOR INT INC.  
 PA  
 PI Bower BS, Fowler T, Phillips JT;  
 PI  
 XX WPI; 1999-395187/33.  
 DR  
 XX Egitr like cellulase  
 PT  
 PS Example; Fig 6; 47pp; English.  
 XX  
 CC The present polypeptide represents a full-length sequence of a  
 CC novel Egitr-like cellulase of Glucoladum roseum. It was deduced  
 CC from a gene sequence isolated from genomic DNA using PCR  
 CC primers (see AAX59180-91) based on conserved motifs (see AAY06325-29)  
 CC of Trichoderma reesei Egitr cellulase and related enzymes. PCR  
 CC has been used to identify novel Egitr-like enzymes, including the  
 CC present protein, from bacterial and fungal sources (see AAY06331-70).  
 CC The sequence shows homology to T. reesei Egitr (see AAY06330). Also  
 CC provided by the invention are vectors, host cells and methods  
 CC for the recombinant production of such enzymes, which can be used  
 CC in the treatment of cellulose-containing textiles, as feed  
 CC additives, in the treatment of wood pulp, in the reduction of  
 CC biomass to glucose, in the stone washing of indigo dyed denim, or  
 CC as laundry detergent components (all claimed).  
 CC  
 SO Sequence 236 AA;

Query Match 92.5%; Score 37; DB 20; Length 236;  
 Best Local Similarity 83.3%; Pred. No. 11e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6  
 KW |||||:  
 DB 63 adsws 68

## RESULT 4

AB14876  
 ID AAB14876 standard; Protein: 236 AA.

AC AAB14876;

DT 21-NOV-2000 (first entry)

DE Glucoladium roseum (3) EGIII-like cellulase.

KM Glucoladium roseum; Trichoderma reesei; endoglucanase III; EGIII;

KW cellulase; mutant; enzyme stability; textile treatment;

OS Glucoladium roseum.

PN WO200037614-A2.

PD 29-JUN-2000.

PF 12-NOV-1999; 99WO-US26704.

PR 18-DEC-1998; 98US-0216295.

PA (GENEV) GENENCOR INT INC.

PI Mitchinson C, Wendt DJ;

DR WPI; 2000-482483/42.

PT Novel endoglucanase III or endoglucanase III-like cellulase useful for

PT treating textiles and wood pulp comprises a substitution or deletion at

PT specified positions in the wild form of endoglucanase III -

PS Example 1; Fig 3; 52pp; English.

CC The present sequence is a cellulase related to endoglucanase III (EGIII)

CC from Trichoderma reesei. EGIII-like genes were isolated from genomic DNA

CC libraries constructed from various microorganisms by PCR. The isolated

CC genes showed significant homology to EGIII from T. reesei. Certain

CC substitution and deletion mutations have been incorporated into EGIII and

CC EGIII-like cellulases to produce variant enzymes with improved stability,

CC e.g. increased resistance to temperature stress. The mutants may be used

CC in textile and wood pulp treatment, as a feed additive, and for reducing

CC biomass to glucose. They are also useful for stonewashing or indigo dyed

CC denim and as an agent in laundry and dish detergents.

CC Sequence 236 AA;

SQ

Query Match 92.5%; Score 37; DB 21; Length 236;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6

DB 63 adsws 68

## RESULT 5

AAV84341

DT 12-JUL-2000 (first entry)

DE Amino acid sequence of an endoglucanase III (EGIII)-like cellulase.

XX Endoglucanase III; EGIII; EGIII-like cellulase; surfactant stability;

KW cellulase; textile processing; textile cleaning; stonewashing;

KW indigo dyed denim; cellulose containing fabric; fabric smoothness;

KW pill removal; fibril removal; cotton; cellulosic fibre; dyeing; detergent;

OS Glucoladium roseum.

PN WO200014208-A1.

PD 16-MAR-2000.

PF 24-AUG-1999; 99WO-US19154.

PR 03-SEP-1998; 98US-0146729.

PA (GENEV) GENENCOR INT INC.

PI Fowler T;

DR WPI; 2000-271052/23.

PT Novel variant endoglucanase III-like cellulases with improved

PT surfactant stability and resistance to temperature stress, useful for

PT textile processing or cleaning, treating wood pulp, food and grain, and

PT reducing biomass to glucose

PS Disclosure; Page 62; 73pp; English.

CC The present sequence represents an endoglucanase III (EGIII)-like

CC cellulase. The cellulase has homology to the Trichoderma reesei EGIII

CC protein. The variant cellulases have improved temperature stability,

CC and improved surfactant stability. The variant cellulases and

CC compositions containing them are used in textile processing or cleaning,

CC e.g. stonewashing of indigo dyed denim, and modifying the texture, feel

CC or appearance of cellulose containing fabrics (e.g. improving fabric

CC smoothness or removing pills and fibrils). The compositions may also be

CC used for the removal of immature or dead cotton from cellulosic fibres

CC or fabric, which can cause uneven dyeing. The cellulase may also be used

CC in a detergent composition for washing laundry and dishes and in the

CC treatment of animal feed, wood pulp, paper, non-animal foods and grains.

CC The enzymes may also be used in the reduction of biomass to glucose.

CC Sequence 236 AA;

SQ

Query Match 92.5%; Score 37; DB 21; Length 236;

Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSWA 6

DB 63 adsws 68

RESULT 6

ABB62635

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 14697.

KW Drosophila; developmental biology; cell signalling; insecticide;

OS Drosophila melanogaster.

PN WO2000171042-A2.

PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PMD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL06738.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 14697; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent.  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AB57737-AB572072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 597 AA;

Query Match 92.5%; Score 37; DB 22; Length 597;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSCWA 6  
 DB 158 sqmswa 163

RESULT 7  
 ID AAU33594  
 AC AAU33594; standard; Protein; 885 AA.  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Pseudomonas aeruginosa cellular proliferation protein #38.  
 XX  
 KW Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX  
 OS Pseudomonas aeruginosa.  
 XX  
 PN WO200170955-A2.  
 PD 27-SEP-2001.  
 PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX

PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS51453.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID NO 5090; 51pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 885 AA;

Query Match 92.5%; Score 37; DB 22; Length 885;  
 Best Local Similarity 83.3%; Pred. No. 4.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSCWA 6  
 DB 563 adwawa 568

RESULT 8  
 ID AAM85714  
 AC AAM85714; standard; Protein; 50 AA.  
 XX  
 DT 07-NOV-2001 (first entry)  
 XX  
 DE Human immune/haematopoietic antigen SEQ ID NO:13307.  
 XX  
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157182-A2.  
 PD 09-AUG-2001.  
 PF 17-JAN-2001; 2001WO-US01354.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 XX

PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214866.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220966.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226661.  
PR 22-AUG-2000; 2000US-0226668.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
DR N-PSDB; AAK58495.  
XX  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX  
PS Claim 11; SEQ ID NO 13307; 3071pp + Sequence Listing; English.  
XX  
XX AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)

CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC XX

SO Sequence 50 AA;

# Query Match

Best Local Similarity 90.0%; Score 36; DB 22; Length 50;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 |||||  
 DB 28 admsw 32

## RESULT 9

AA65780 standard; Protein; 174 AA.

AA65780;

27-MAR-2001 (first entry)

Cysteine protease #21.

Cell death modulator; programmed cell death; PCD; apoptosis;  
 forestry plant.

Eucalyptus grandis.

MO200075331-A1.

14-DEC-2000.

02-JUN-2000; 2000MO-NZ00086.

04-JUN-1999; 99US-0325932.

(GENE-) GENESIS RES & DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Flinn B, Lasham A;

WPI: 2001-061724/07.

N-PSDB; AAF44807.

Novel defender against cell death polynucleotide useful for modulating  
 programmed cell death pathway and specific development pathways in  
 forestry plant.

Claim 22; Page 116; 142pp; English.

The present invention relates to coding sequences (see AAF44740-F44840  
 CC and AAF44843-F44844) and proteins (see AAB65714-B65814) involved in  
 CC programmed cell death (PCD; apoptosis). The coding sequences and proteins  
 CC of the present invention are useful for modulating a PCD or cell death  
 CC pathway and various developmental pathways in a forestry plant, by  
 CC stably incorporating one of the present coding sequences into the genome  
 CC of the forestry plant, where the coding sequence provides a PCD pathway  
 CC that is not present in a native form of the forestry plant.

SO Sequence 174 AA;

# Query Match

Best Local Similarity 90.0%; Score 36; DB 22; Length 174;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ADMSW 5  
 |||||  
 DB 109 admsw 113

## RESULT 10

AA30839 standard; Protein; 218 AA.

AA30839;

13-MAR-1998 (first entry)

Myokl protein myoklp.

Myotonin kinase; human; myokl; myoklreg; myoklreg; diagnosis;  
 prevention; treatment; myotonic dystrophy.

Homo sapiens.

CA2153706-A.

21-MAY-1997.

20-NOV-1995; 95CA-2153706.

20-NOV-1995; 95CA-2153706.

(PRED/) PREDDIE R E.

Bergmann JE, Preddie RE;

WPI: 1997-403120/38.

N-PSDB; AAT91960.

New isolated myoklreg and myokl nucleic acid sequences - used to  
 develop products for the pre-symptomatic diagnosis, prevention and  
 treatment of myotonic dystrophy

Claim 2; Page 7-8; 18pp; English.

The present sequence represents the new isolated myokl protein, myoklp.  
 CC Myokl is located within the 3' UTR of the human myotonin kinase (mk)  
 CC gene on the complementary DNA strand. A part of the myokl coding region  
 CC is complementary to the region of the mk gene containing the (CTG)<sub>n</sub>  
 CC repeat believed to be the genetic cause of myotonic dystrophy (MD). The  
 CC Myoklp protein shows homology to the tachykinin-related protein family,  
 CC e.g. the androgen receptor and Ataxin 1, however the protein also shows  
 CC features evolutionarily related to the adrenocorticotrophic hormone  
 CC (ACTH). Inhibitors of Myokl and Myoklp, e.g. antisense sequences and  
 CC antibodies, are used for the pre-symptomatic diagnosis, prevention and  
 CC treatment of myotonic dystrophy (MD). Myokl and Myoklp can be used for  
 CC developing transgenic systems in which models of MD can be produced and  
 CC studied. From a clinical aspect, because the protein is expressed only  
 CC in response to specific stressful events and is involved in causing  
 CC shock in humans, it is possible, without serious side effects, to block  
 CC the transcription/expression of Myokl mRNA or activity of Myoklp in  
 CC people presymptomatic for MD and stop or prevent the clinical symptoms  
 CC of the disease.  
 CC N.B. The new isolated gene is specified as myokl in the claims section,  
 CC but is referred to as myokl in the disclosure.

Sequence 218 AA;

# Query Match

90.0%; Score 36; DB 18; Length 218;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DMSMA 6  
DB 49 dmswa 53

## RESULT 11

AAB65779 ID AAB65779 standard; Protein; 225 AA.

XX AC AAB65779;

DT 27-MAR-2001 (first entry)

DE Cysteine protease #20.

XX KM Cell death modulator; programmed cell death; PCD; apoptosis;  
foresty plant.

XX OS Eucalyptus grandis.

XX PN WO200075331-A1.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-N200086.

XX PR 04-JUN-1999; 99US-0325932.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX PI Flinn B, Lasham A;

XX DR WPI; 2001-061724/07.

XX N-PSDB; AAF44806.

PT Novel defender against cell death polynucleotide useful for modulating  
PT Programmed cell death pathway and specific development pathways in  
PT foresty plant -

XX PS Claim 22: Pages 115-116; 142pp; English.

CC The present invention relates to coding sequences (see AAF44740-F44840  
CC and AAF44843-F44844) and proteins (see AAB65714-B65814) involved in  
CC programmed cell death (PCD; apoptosis). The coding sequences and proteins  
CC of the present invention are useful for modulating a PCD or cell death  
CC pathway and various developmental pathways in a foresty plant, by  
CC stably incorporating one of the present coding sequences into the genome  
CC of the foresty plant, where the coding sequence provides a PCD pathway  
CC that is not present in a native form of the foresty plant.

XX SQ Sequence 225 AA;

Query Match 90.0%; Score 36; DB 22; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSM 5  
DB 100 admsw 104

## RESULT 12

AAM39444 ID AAM39444 standard; Protein; 276 AA.

XX AC AAM39444;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2589.

XX KM Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX KM peripheral nervous system; neuropathy; central nervous system; CNS;  
XX KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
XX KM chemokine; thrombolytic; drug screening; arthritis; inflammation;  
XX KM leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX N-PSDB; AAI58600.

PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -

XX PS Example 4; SEQ ID NO 2589; 10078pp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic, and  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC Specification.

XX SQ Sequence 276 AA;

Query Match 90.0%; Score 36; DB 22; Length 276;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSM 5  
DB 162 admsw 166

## RESULT 13

AAB42053 ID AAB42053 standard; Protein; 277 AA.

XX AAB42053;  
 AC  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1817 polypeptide sequence SEQ ID NO:3634.  
 XX  
 KM Human; open reading frame: ORFX; detection: cytostatic; hepatotropic;  
 KM vulnary; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
 KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KM hypotensive; dermatological; immunosuppressive; antineoplastic;  
 KM antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;  
 KM antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KM neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KM cholesterol ester storage; systemic lupus erythematosus; infection;  
 KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KM bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KM thrombosis; contraceptive.  
 KM  
 OS Homo sapiens.  
 XX  
 PN MO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000MO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkens RA, Leach M;  
 FI  
 DR WPI: 2000-602362/57.  
 DR N-PSDB: AAC76262.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 PS  
 PS Claim 11; Page 2792-2793; 5507pp; English.

Query Match 90.0%; Score 36; DB 21; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ADMSW 5  
 Db 163 admsw 167  
 RESULT 14  
 ABG24678  
 ID ABG24678 standard; Protein: 289 AA.  
 XX  
 AC ABG24678;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #24669.  
 XX  
 KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KM food supplement; medical imaging; diagnostic; genetic disorder.  
 KM  
 OS Homo sapiens.  
 XX  
 PN MO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001MO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HSE-) HXSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 FI  
 DR WPI: 2001-639362/73.  
 DR N-PSDB: AAS88865.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 PS  
 PS Claim 20; SEQ ID NO 55037; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 289 AA;



Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
11111  
DB 283 admsw 287

RESULT 15  
AAM41230  
ID AAM41230 standard; Protein: 371 AA.  
XX  
AC AAM41230;  
XX  
DT 22-OCT-2001 (first entry)  
XX

DE Human polypeptide SEQ ID NO 6161.

XX Human; nootropic; immunosuppressant; cyostatic; gene therapy; cancer;  
KM peripheral nervous system; neuropathy; central nervous system; CNS;  
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;  
KM Chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukemia.  
XX  
XX

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PE 26-DEC-2000; 2000WO-US34263.

XX  
PR 21-JAN-2000; 2000US-0486725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX

DR WPI: 2001-442253/47.  
DR N-PSDB; AAI60386.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX

PS Example 2; SEQ ID NO 6161; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX

SQL Sequence 371 AA;

Query Match 90.0%; Score 36; DB 22; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ADMSW 5  
11111  
DB 257 admsw 261

Search completed: September 3, 2002, 11:51:20  
Job time: 191 sec